

Proper Orthogonal Decomposition for Surrogate Models in Aerodynamics

Original

Proper Orthogonal Decomposition for Surrogate Models in Aerodynamics / Dolci, Valentina. - (2017).
[10.6092/polito/porto/2678186]

Availability:

This version is available at: 11583/2678186 since: 2017-08-10T11:48:43Z

Publisher:

Politecnico di Torino

Published

DOI:10.6092/polito/porto/2678186

Terms of use:

Altro tipo di accesso

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

POLITECNICO DI TORINO



Department of Mechanical and Aerospace Engineering

Ph.D. in Fluid Dynamics

Proper Orthogonal Decomposition for Surrogate Models in Aerodynamics

Supervisor

Prof. Renzo Arina

Student

Valentina Dolci

August 2017, XXVIII ciclo, ING-IND/06

Abstract

This study describes the design and implementation of surrogate models for aerodynamic optimization or database generations. Two different methods are presented: the first one follows the classical methodology: a parametric POD is applied to a set of initial solutions or snapshots obtained with an high fidelity CFD model. With respect to approaches presented in literature, in this research work no truncation of the POD modes is performed and they are all used to construct the surrogate model. Several applications are presented: a backward facing step case, the analysis of the flow around a NACA 0012 airfoil and a RAE 2822 supercritical airfoil, the optimization of an automotive external shape and a database generation of a three-dimensional aircraft.

The second methodology presented for the construction of a surrogate model is a novel technique in which POD is applied directly to the parameters in order to obtain a representation of the problem in a reduced-order space. An optimization is performed in this new space and the optimum point is then re-projected into the original framework. The methodology is applied and tested in the minimization of the Viermin function and the POD surrogate model shows good agreements with the analytical solution.

Nomenclature

(1)	point with all factors at low level
+	high level
−	low level
α	angle of attack (AOA)
β	vector of unknown model parameters
$\varphi(x)$	Proper orthogonal decomposition (POD) basis
\mathbf{X}	sampling plane
\mathbf{x}	vector of design variables
δ	parameter combination set used for the PODI (proper orthogonal decomposition with interpolation)
ϵ	random error with zero mean
λ	POD eigenvalue
ϕ	radial function
σ	shape parameter of a radial basis function

σ^2	variance
θ	correlation parameter
\tilde{A}_r^i	pseudoinverse of matrix \tilde{A}_r
\tilde{A}_r	reduction matrix
\tilde{f}	surrogate model approximation of the real function \mathbf{f}
\tilde{p}^*	minimum value of the function in the reduced dimension space
A	interpolation matrix
a	parameter combination with high level of factor x_1
ab	combination of factors x_1 and x_2 at high level
b	parameter combination with high level of factor x_2
$B(x)$	trend function of the Kriging model
c	parameter combination with high level of factor x_3
c_D	drag coefficient
C_j	POD coefficients
c_L	lift coefficient
COV	covariance matrix
D	design space
e	POD eigenvector
$E\%$	L_2 percentage error

$e\%$	maximum relative error
E_n	normalized root mean square error
e_{rel}	relative error
f	function, or output of the problem
$f_v(\mathbf{x})$	Viermin function
$g(\mathbf{x})$	low order polynomial
k	cardinality of the design space
l	level of the grid
M	number of snapshots
n	number of samples
p	fraction of the full factorial design
$p(x)$	polynomial
p^*	optimum value of the function $f(\mathbf{x})$
R	correlation matrix
r	number of regressors
$s(\mathbf{x})$	Kriging predictor
V	state matrix
w	radial basis function coefficient
x_1	factor of a design screening plane

x_2	factor of a design screening plane
x_3	factor of a design screening plane
y	output of the problem
y_1	width of the trunk
$Z(x)$	Gaussian distribution of the Kriging model
z_1	height of the trunk
z_2	height of the diffuser

Contents

1	Introduction	1
1.1	Main Features	4
1.2	Sampling of the Design Space	5
1.2.1	Elementary Effects	8
1.2.2	Factorial Design	9
1.2.3	Central Composite Design	13
1.2.4	Fractional Factorial Design	14
1.2.5	Quadtree Distribution	19
1.2.6	Latin Hypercube Sampling	20
2	Surrogate Model	22
2.1	Kriging	23
2.2	Proper orthogonal decomposition	24
2.2.1	Response Surface Methodology	30
2.3	Proper Orthogonal Decomposition applied to the Parameter Space . .	34
2.3.1	Methodology	35
3	Applications	38
3.1	Backward Facing Step	39
3.1.1	Problem setting	39

CONTENTS

3.1.2	POD reconstruction	41
3.1.3	POD reconstruction with different snapshot sets	44
3.2	NACA 0012 Surrogate Model	48
3.2.1	Interpolation technique - preliminary phase	48
3.2.2	Influence of the snapshot number	52
3.2.3	Influence of the snapshot position - Quadtree initial distribution	56
3.3	RAE 2822 Surrogate Model	60
3.3.1	Problem Setting	60
3.3.2	Results	61
3.4	Drag Coefficient Optimization of an Automotive Shape	68
3.4.1	Problem and CFD setting	68
3.4.2	Optimization	69
3.5	3D Aircraft Surrogate Model	75
3.5.1	Problem Setting	75
3.5.2	Results	76
3.6	Optimization of the three-parameter Viermin function	79
4	Conclusions	82

Chapter 1

Introduction

“Simplicity is the final achievement. After one has played a vast quantity of notes and more notes, it is simplicity that emerges as the crowning reward of art.”

These words from the polish composer Frédéric Chopin represent the tendency of the humankind to schematize the reality and remember only the principal features of the surrounding nature.

The topic of the present work will be the construction of surrogate models for aerodynamic applications. A surrogate model in fact is a way to extract the main components of a particular discretization of the reality in order to make the data storage and the comprehension of a specific problem physics easier.

The increasing complexity of engineering systems has caused an increased computational burden to simulate how a complex product will perform. In the aerospace industry for example is possible to simulate the entire flow around a generic HSCT (High Speed Civil Transport) configuration, the vibration and noise experienced inside a tilt rotor engine bay and an airplane life when subjected to repeated stressful conditions. Moreover this simulations can be quickly modified to reflect changes and optimization loops can be used to improve engineering design.

However any optimization algorithm applied directly to the full model will be slow due to the long running time that the simulations can take. Even if one would invest all the computational budget in applying an optimization algorithm directly to the simulations, this unlikely will be enough because rarely a single optimization result settles any design issue.

For example, if the result is not satisfactory, parameter sweeps, plots of input-output relationships, optimization with different starting point and constraints may be required and all this, of course, implies more simulations. On the other hand, if the result is satisfactory, further investigations will be necessary and again, this involves doing more simulations. If all the available resources were exhausted solving the first optimization problem, all these following analyses would not be possible, or at least lead to missed deadlines.

The basic idea in the surrogate model approach is to consume the computational budget not in a single optimization loop but in developing fast approximations to the full-model computer codes.

The general approach of a surrogate model is the construction of a simplified mathematical representation of the computationally expensive complete description of the problem (i.e. in the aerodynamic case the Navier-Stokes equations). The new model can then be used in simple or multidisciplinary design optimization, design space exploration, reliability analysis and in every field where many simulations are required.

Since the approximation model acts as a “model of the model”, it can be called metamodel. A variety of surrogate models can be constructed (e.g., polynomial response surfaces, kriging models, radial basis functions, neural networks, multivariate adaptive regression splines), and recent reviews and comparisons can be found in [11].

In aerodynamic problems, the main way to acquire data is to reproduce the flow

fields around the specific geometry using computational fluid dynamics (CFD) techniques. These techniques use numerical methods to solve the already mentioned Navier-Stokes equations and can be considered the high-fidelity or full model of the problem in fact if compared with experimental methods, they are reliable. The major drawback is that they are computationally expensive, highly memory demanding, and time consuming. Simplified forms of the governing equations can be used but the increasing processing speeds and memory capacities of modern computers cause simultaneously the demand for more accurate real calculations and more complicated designs.

Similar results can be obtained in a flexible, and cheaper way: aerodynamic design optimization makes nowadays a great use of surrogate models in order to handle sophisticated geometries, simulate more realistic flight conditions, and satisfy increasing design requirements.

Such models can be used to bridge between the different levels of sophistication that characterize fidelity physics based simulation codes, or between predictions and experiments. They seek to fulfill the gaps between the limited analysis runs of the full models that can be afforded with the necessarily limited computing power.

Another important problem that arises using a direct interface between optimization methods and high fidelity analysis codes is the occurrence of numerical noise.

High fidelity model numerical noise is principally caused by the discrete representation of continuous governing equations and physical quantities. It can include numerical round-off errors and can appear if the iterative convergence of calculations is not reached. These errors can lead the optimization algorithm to wrong directions and cause local search methods to stuck in local optima. The adoption of a surrogate model can avoid this complication.

1.1 Main Features

The surrogate model approach is based on the constraint that its adoption will make possible to save many orders of magnitude of time with respect to the primary source while still generating an acceptable error when used away from known data points. Two factors have to be considered therefore: significant speed increase and sufficient accuracy. During the building of a surrogate model these requirements will be often conflicting and the designer will have the difficult task to balance them.

The approximation model has to be as accurate as possible over the complete domain of interest while minimizing the simulation cost. This constraint implies multiple sub-problems: together with the selection of the model type and its complexity, where to collect the initial data and how to estimate the model quality should be decided.

The data collection is fundamental since it is the most computationally expensive aspect and the optimal data distribution is not known a priori.

Though a totally fixed approach won't be appropriate since the data itself may well influence the directions taken, the main steps to be performed when building a surrogate model can be summarized below:

- Sampling of the design space,
- Selection of a surrogate model form,
- Model validation.

Firstly, an initial data set is obtained by sampling the design decision space. During this step the expensive high-fidelity analysis code is used. The initial set will contain relating inputs and outputs and different candidate designs are generated and analysed. The position and quantity of the training set points will be one of

the central aspects for the accuracy of the surrogate model because the amount and location of the initial information supplied to the surrogate will define its capability to generate predictions at unknown points. During the data building therefore two conflicting targets have to be taken into account: increasing the model accuracy by incorporating as large training set as possible and decreasing the computational cost of data generation by selecting as few training points as possible.

After this first phase of design space sampling, an appropriate surrogate form must be selected and applied to the available data. The parameters of the model must be estimated and the model must be assessed for accuracy. A distinction between parametric or nonparametric surrogate models can be made. The first class of models is simpler and it would take less effort to be implemented. On the other hand parametric models are recommended when applied to cases with elaborated input-output relations.

After the selection of the surrogate model form, an updating of the model can be foreseen. To update the model new information can be supplied through the so called infill points. In other words a new sampling of the parameter space will be generated adding new points to the initial data set. The infill points can be located either in the areas where the surrogate presents the higher prediction errors or, alternatively, in the zones detected by the surrogate where particularly interesting combinations of design variables can lie.

1.2 Sampling of the Design Space

Let's consider a continuous quantity $f(\mathbf{x})$, function of a particular vector $\mathbf{x} \in D \subset \mathbb{R}^k$. In the framework of surrogate model techniques and optimization in general, the vector \mathbf{x} is called vector of design variables and D is defined as the design space or design domain.

If we are interested in analyse f , we can do it through discrete observations or samples $f(\mathbf{x}^{(i)})$. In order to get this samples, the high fidelity model has to be used and to keep the computational and time effort low, the observation set has to be as sparse as possible. A sampling plan \mathbf{X} can be defined as $\mathbf{X} = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}\}$. The surrogate model is able to use this sample plane to build an approximation of $f(\mathbf{x})$, that we can call $\tilde{f}(\mathbf{x})$, which can then be used to make cheaper performance predictions for any vector $\mathbf{x} \in D$.

The features of the sampling plane \mathbf{X} will be crucial in defining the accuracy of the surrogate model predictions of unseen data. For example, a point concentration at the extreme parameter values of a design space may leave a great lack of information in the centre. In the same way, collocate points exclusively in internal portions of the sampling plane, causing far-reaching extrapolations from the surrogate model in the external parts of the domain, may lead to wrong global conclusions based on irregular, local knowledge of the objective behaviour.

During the design of a sampling plane for a surrogate model implementation, one of the principal problems of this kind of approach arises. Let's establish that we built a surrogate model in a one-dimensional parameter space and that we are satisfied with its prediction accuracy level. The total number of samples used for this task is equal to n . If the same accuracy level is needed in a k -dimensional space, with $k > 1$, the number of observations that will be necessary is equal to kn . This problem is often referred to as 'the curse of dimensionality' and from the previous formula we can easily see how massive is the impact of the design variable number on the amount of required experiments.

Another important feature that has to be taken into account during the design of a computer experiment is that the sources of error of physical and computational experiments are different. During a physical experiment in fact, human, systematic and random errors can be present. The systematic error is the adding of a constant

bias to the result. The random error on the other hand is caused by unknown and unpredictable changes in the measuring instruments or in the environmental conditions.

In a computational experiment the first two error sources will still be present but this kind of experiment is deterministic and therefore not affected by random errors.

With that in mind is clear that the screening design techniques have to be different between physical and computed experiments. For example the concept of replication, used in physical experiments to weaken noise in the results, is pointless in the computed experiment field because replying a simulation with the same initial conditions will lead to the exact equal results. The same way of thinking can be applied to the concepts of blocking and randomization, used in the experimental field to contrast random errors.

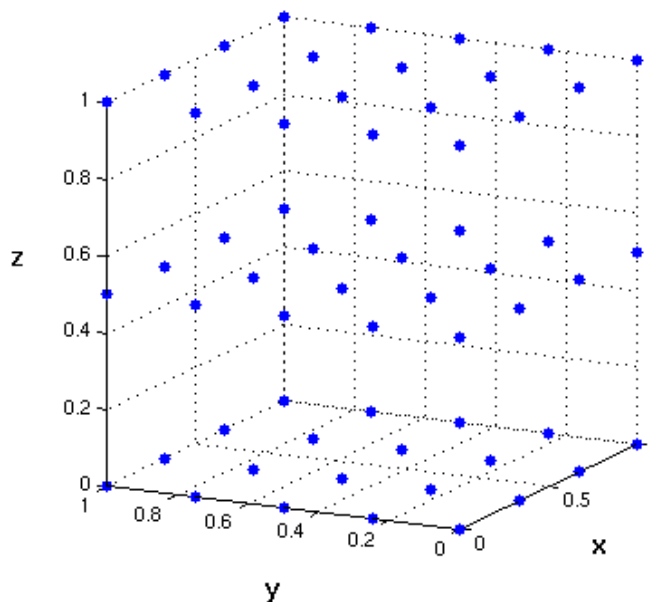


Figure 1-1: *Visualization of the distribution of points in a full factorial plane*

In general, a surrogate model has a greater accuracy nearby the points belonging to the sampling set. If a uniform level of accuracy within the design space is required, the sampling plane will have to be characterized by a uniform location of the points.

This kind of sampling plan is called space-filling and the easiest way to achieve this property is to use the full factorial sampling technique, visualized in fig. 1-1. In full factorial designs, an objective function evaluation is performed at every combination of the factor levels, were with the term factors we are referring to the design variables of the objective function $f(\mathbf{x})$.

1.2.1 Elementary Effects

As we saw in the previous sections, the behaviour of the objective function $f(\mathbf{x})$ has to be inspected through a computational experiment, in which a number of evaluations are made and an a-posteriori analysis of the results is performed. Each evaluation corresponds to a run of the high fidelity model at different input values.

Considering the 'course of dimensionality' problem, that strongly depends on the number of variables, in the initial phases of developing a surrogate model, an important experimental activity is the identification of which inputs have influence on the outputs and cannot reasonably be ignored. This process is often referred to as 'Factor screening' or 'Sensitivity analysis'.

If at least the first derivative of $f(x)$ with respect to each variable exists, the value

$$\partial_i(\mathbf{x}) = \left. \frac{\partial f}{\partial x_i} \right|_{\mathbf{x}} \quad (1.1)$$

can be considered as an indication of the influence of x_i on $f(x)$.

In particular if

- $\partial_i(\mathbf{x})$ is zero over all values of \mathbf{x} , we can say that the effect of \mathbf{x}_i is negligible

on $f(\mathbf{x})$;

- if $\partial_i(\mathbf{x})$ is a non-zero constant over all values of \mathbf{x} , the effect is linear or additive;
- if $\partial_i(\mathbf{x})$ is a function of \mathbf{x}_i , the effect is nonlinear and
- if $\partial_i(\mathbf{x})$ is a function of more than one \mathbf{x}_j , with $j \neq i$, the effect is involved in interactions with other inputs.

Estimates of mean and standard deviations of these distributions can be considered to distinguish the important inputs from the negligible ones. A general definition of an elementary effect, that is the change in an output due solely to changes in a particular input, can be found in [17]. Considering a regular grid, $x_i \in \{0, (l-1)/2, \dots, 1\}$ for $i = 1, \dots, k$, where l is the level of the grid, the elementary effect of x_i can be written as

$$d_i(\mathbf{x}) = \frac{f(x_1, x_2, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_k) - f(\mathbf{x})}{\Delta}, \quad (1.2)$$

where the divisor Δ is a fixed step size and a predetermined multiple of $1/(l-1)$. A preliminary screening has to be foreseen to estimate the elementary effects keeping in mind that when strong nonlinearities or interactions exist, there is the possibility that a design of experiment reduced with respect to the number of inputs can not generate sufficient information.

1.2.2 Factorial Design

Together with elementary effects we can define the joint effects [19] as main effects and interactions. In order to investigate the joint effects of several variables, factorial designs are widely used in computer experiments. The simpler factorial plane that can be considered is the one built imposing only two levels at each variable. It is named two-level full factorial design and indicated as 2^k . This design is a basic building

block that can be used to create more complex planes. Let's consider 2 variables, x_1 and x_2 . To build the screening plane each variable will assume a high and a low value, that will be the levels of the factorial plane. They can be associated to the signs $+$ and $-$. The high level of each factor can be indicated with lowercase letters. In fig. 1-2, the plane is represented. The point (1) corresponds to the low level of both factors. a stays for the combination: factor x_1 at high level and factor x_2 at low level and so on.

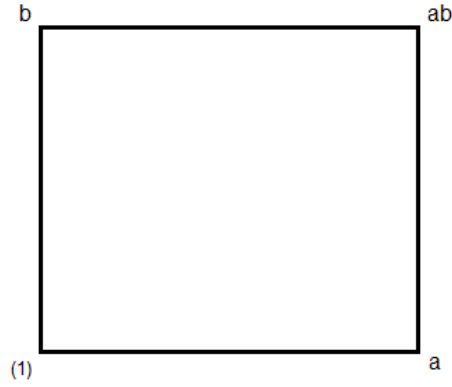


Figure 1-2: *Graphic visualization of a 2-level full factorial plane for two factors*

The main or average effect of a factor can be easily computed from this kind of factor planes as

$$E(x_1) = \bar{f}_{x_1^+} - \bar{f}_{x_1^-} = \frac{f(ab) + f(a)}{2} - \frac{f(b) + f((1))}{2}, \quad (1.3)$$

where with $\bar{f}_{x_1^+}$ is indicated the average response of the output f , when the variable x_1 is at high level and with $\bar{f}_{x_1^-}$ the average response of the output when x_1 is at low level.

In the same way the main effect of x_2 can be expressed by

$$E(x_2) = \bar{f}_{x_2^+} - \bar{f}_{x_2^-} = \frac{f(ab) + f(b)}{2} - \frac{f(a) + f((1))}{2}. \quad (1.4)$$

The interaction effect IE is equal to

$$IE(x_1, x_2) = \frac{f(ab) + f(1)}{2} - \frac{f(a) + f(b)}{2}. \quad (1.5)$$

Considering fig. 1-2 again, the main effects of the output can be computed combining the sides of the square and the interaction effect corresponds to a combination of the diagonals. The magnitude and direction of these effects can be used to understand importance and action of a factor on the output. If the main effect of x_1 is positive, for example, the increasing of x_1 will correspond to an increasing of the output. On the other hand, if the main effect of a factor is small with respect to the other effects, the corresponding factor can be neglected in the consequent analyses and in the building of the surrogate model.

If we are considering three factors, x_1 , x_2 and x_3 , the corresponding simpler full factorial plane is the 2-level 3-factor plane, indicated as 2^3 . This design is composed by eight runs, listed in tab. 1.1 using the geometric notation with $+$ and $-$ to indicate high and low level of a variable. The combination can be represented graphically as a cube, as shown in fig. 1-3

To estimate the main effects of the three factors, the sides of the faces of the cube can be combined, in a similar way to the 2^2 factorial plane. For example the main effect of x_1 is computed as

$$E(x_1) = \bar{f}_{x_1^+} - \bar{f}_{x_1^-} = \frac{f(a) + f(ab) + f(abc) + f(ac)}{4} - \frac{f(b) + f((1)) + f(c) + f(bc)}{4} \quad (1.6)$$

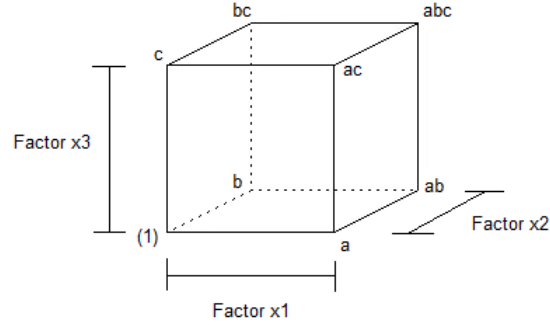


Figure 1-3: *Graphic visualization of a 2^3 full factorial plane*

and the interaction IE between x_1 and x_2

$$IE(x_1, x_2) = \frac{f(ab) + f(1) + f(c) + f(abc)}{4} - \frac{f(a) + f(b) + f(bc) + f(ac)}{4} \quad (1.7)$$

is expressed considering the diagonal planes. Due to this reason the use of 2^k full factorial designs for preliminary screening of the parameter space allows to perform a minimum number of runs avoiding redundancy and to have an estimate of main and interaction effects between the parameters. Moreover this kind of planes has the

Table 1.1: Factor combinations of a 2^3 full factorial plane

Run	x_1	x_2	x_3
1	—	—	—
2	+	—	—
3	—	+	—
4	+	+	—
5	—	—	+
6	+	—	+
7	—	+	+
8	+	+	+

important property of orthogonality: the main effect of a factor can be computed independently from the other variables and the same happens for the interactions.

A general 2^k design will contain k main effects, $\binom{k}{2}$ two-factor interactions, $\binom{k}{3}$ three-factor interactions and so on and one k -factor interaction. In total, it is possible to estimate $2^k - 1$ effects.

Full factorial designs are very useful in estimating the effects of the variables, but they are affected by the curse of dimensionality. If the number of factors k is very high, the runs composing the design will grow exponentially. In that case an a-priori knowledge can come to hand: the sparsity of effects principle. According to this theory, if we are dealing with a high number of variables, only some main effects and low-order interactions will be dominant, with negligible high-order effects.

1.2.3 Central Composite Design

Another process that can be useful to improve the accuracy of the joint-effect estimations using a 2^k factorial plane is the addition of the center point to the design. A two-level factorial design, in fact, is typically adopted to build a first-order model of the problem. In [19] the addition of replicates of the center point is recommended for experimental designs. Even in the case of computer designs, adding the center point (obviously avoiding replications) can provide information if some kind of curvature, and therefore a second-order model, is more suitable for the problem. Typically the center point is available as base condition of the simulation and there is no additional cost in considering it in the parameter space sampling. If we are using a 2^2 design for example, a comparison can be made between the average response of the four combinations on the vertices of the square (fig. 1-2) and the response of the central point. If this two values are similar, a first-order model can be accurate to describe the problem. If not, a more complex model should be adopted.

If the fitting of a second-order model is needed, a two-level full factorial design cannot be adopted but we can extend this kind of design adding axial runs. In this way a so-called central composite design is generated. In particular, $2k$ axial runs should be added.

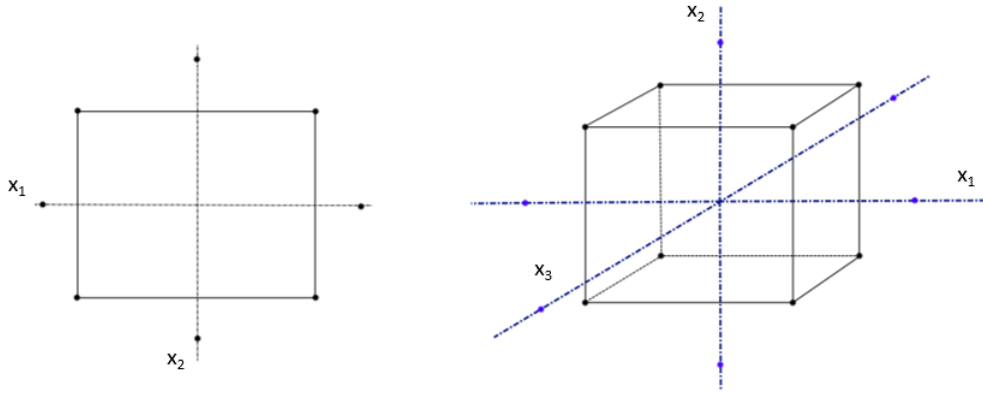


Figure 1-4: *Central composite design for two factors (left) and three factors (right)*

1.2.4 Fractional Factorial Design

A way to partially reduce the number of runs needed to create a sampling plane, is the adoption of fractional factorial designs. In this way main effects and interactions can be computed using only a fraction of the original design and certain high order estimations will be neglected.

The easier type of fractional plane is the one-half fraction of a two-level design. Consider for example a 2^3 run. The corresponding one-half fractional factorial design is composed by $8/2$ runs and indicated as 2^{3-1} . Tab. 1.2 describes the combinations used to calculate main effects and interactions of the three factors x_1 , x_2 , x_3 . The

Table 1.2: Factor combinations of the two 2^{3-1} fractional factorial planes

Run	I	$E(x_1)$	$E(x_2)$	$E(x_3)$	$IE(x_1, x_2)$	$IE(x_1, x_3)$	$IE(x_2, x_3)$	$IE(x_1, x_2, x_3)$
a	+	+	–	–	–	–	+	+
b	+	–	+	–	–	+	–	+
c	+	–	–	+	+	–	–	+
abc	+	+	+	+	+	+	+	+
ab	+	+	+	–	+	–	–	–
ac	+	+	–	+	–	+	–	–
bc	+	–	+	+	–	–	+	–
(1)	+	–	–	–	+	+	+	–

column I corresponds to the identity column. To obtain the interactions, the corresponding columns of the main effects can be multiplied. The two complementary fractional factorial designs are shown in the top and bottom part of the table. The first plane, composed by the first four runs, is built using all the plus-sign combinations of the last column $IE(x_1, x_2, x_3)$. This column is called the generator. The alternate plane, containing the last four runs, is composed by the combinations corresponding to a minus sign of the interaction $IE(x_1, x_2, x_3)$. The two complementary designs can be observed in fig. 1-5. From this picture an important property of fractional factorial designs can be clear: the runs composing these designs in fact can be combined in order to create larger planes to estimate factor effects and interactions.

In a 2^{3-1} design, some main effects are confused or aliased with two-factor interactions. For example

$$E(x_1) = \frac{f(a) - f(b) - f(c) + f(abc)}{2} = E(x_2, x_3), \quad (1.8)$$

$$E(x_2) = \frac{-f(a) + f(b) - f(c) + f(abc)}{2} = E(x_1, x_3) \quad (1.9)$$

and so on. If main effects are aliased with two-factor interactions, the corresponding

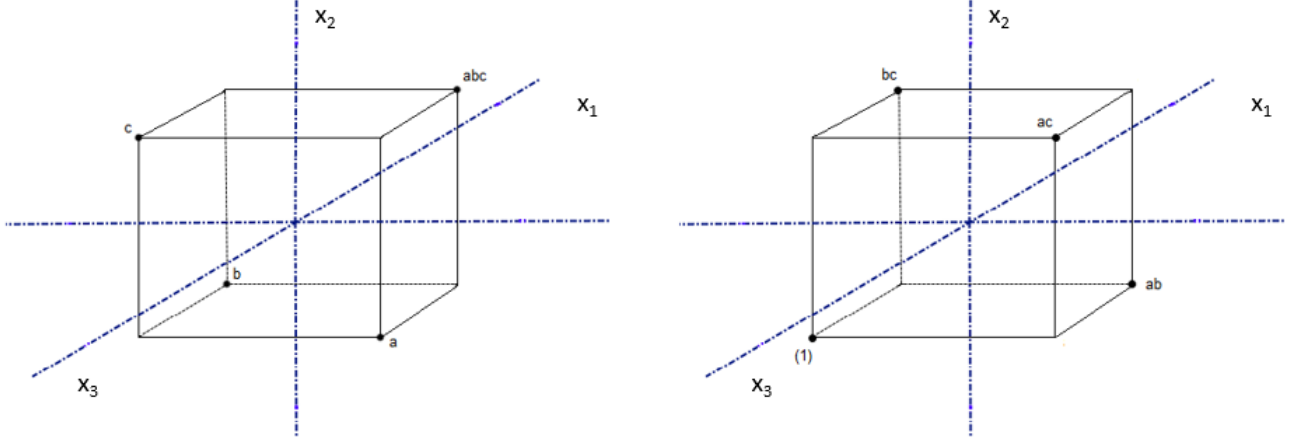


Figure 1-5: *Fractional factorial planes 2^{3-1}*

plane can be called a resolution III design. Using this design, main effects can be estimated without alias.

In a resolution IV design, on the other hand, two-factor interactions are aliased with each other and in a resolution V design two-factor interactions are aliased with three-factor interactions. Of course a high resolution is preferable, keeping at the same time a low number of runs.

An example of a resolution IV design is the 2^{4-1} fractional factorial design. This design is built considering four factors x_1 , x_2 , x_3 and x_4 and listed in tab. 1.3. A graphical representation of a resolution V design can be found in fig. 1-6 where the 2^{5-1}_V is shown.

Another important property of fractional factorial designs is that, considering a design of resolution r , it contains the corresponding full design in any subset of $r - 1$ factors and can be projected onto this subspaces as can be seen in fig. 1-7. If we are dealing with a larger number of factors, smaller fractions of the full factorial design can be employed. For example a one-quarter fraction of a 2^k design will contain 2^{k-2}

Table 1.3: 2_{IV}^{4-1} fractional factorial design

Run	$E(x_1)$	$E(x_2)$	$E(x_3)$	$x_1 \cdot x_2 \cdot x_3$	Treatment combination
1	—	—	—	—	(1)
2	+	—	—	+	ad
3	—	+	—	+	bd
4	+	+	—	—	ab
5	—	—	+	+	cd
6	+	—	+	—	ac
7	—	+	+	—	bc
8	+	+	+	+	abcd

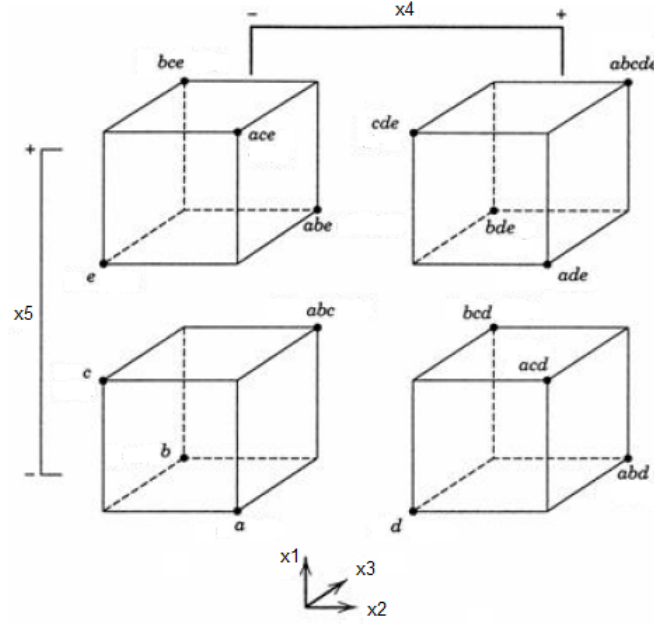


Figure 1-6: 2_V^{5-1} fractional factorial design

runs. The resolution of these designs will be lower and, if the 2^{k-1} design was built using a single generator, now two generators should be used. A 2_{IV}^{6-2} is presented in tab. 1.4. In general a 2^{k-p} has p generators and each effect has 2^{p-1} aliases.

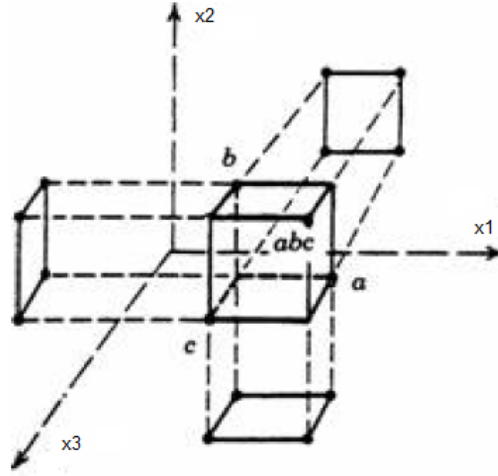


Figure 1-7: *Projection of a 2^{3-1}_{III} fractional factorial design on three 2^3 full factorial designs*

Table 1.4: 2^{6-2}_{IV} fractional factorial design

Run	$E(x_1)$	$E(x_2)$	$E(x_3)$	$E(x_4)$	$x_1 \cdot x_2 \cdot x_3$	$x_2 \cdot x_3 \cdot x_4$
1	—	—	—	—	—	—
2	+	—	—	—	+	—
3	—	+	—	—	+	+
4	+	+	—	—	—	+
5	—	—	+	—	+	+
6	+	—	+	—	—	+
7	—	+	+	—	—	—
8	+	+	+	—	+	—
9	—	—	—	+	—	+
10	+	—	—	+	+	+
11	—	+	—	+	+	—
12	+	+	—	+	—	—
13	—	—	+	+	+	—
14	+	—	+	+	—	—
15	—	+	+	+	—	+
16	+	+	+	+	+	+

1.2.5 Quadtree Distribution

In the previous sections we described factorial planes and central composite design approach. Many sampling distributions have been proposed in literature and in this section the quadtree distribution is briefly described.

Quadtree is a specific subdivision of 2-dimensional spaces used first in digital imaging. It can be considered as a hierarchical data structure [3]. The subdivision is described by the so called tree data structure that is an iterative splitting of the area of interest. At the beginning we can consider the whole space as a cell. This cell has four vertices at its corners. In the first iteration the single cell is divided into four sub-cells and in the following iterations each cell can be divided again in other four sub-cells.

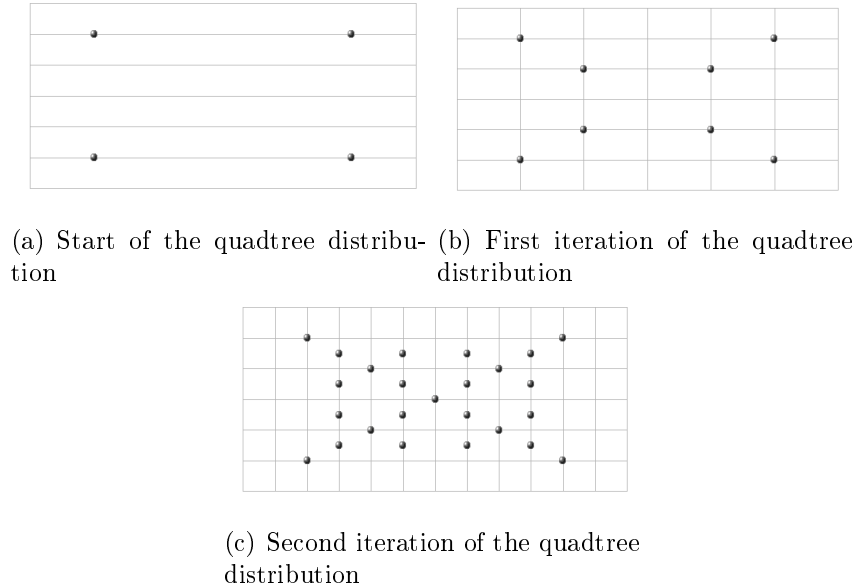


Figure 1-8: Iterations of a quadtree distribution

In Fig. 1-8 the quadtree approach is visually explained. Examples of a quadtree distribution applied to the initial sampling for the building of a surrogate model can

be found in [8].

1.2.6 Latin Hypercube Sampling

The previous ways of sampling the parameter space adopted a uniform sampling, as in the case of factorial planes or central composite designs, or a structured sampling, and that is the case of quadtree distribution. The Latin hypercube sampling is still a stratified sampling but it belongs to the class of random samplings.

We call Latin square a grid with a different value in each row and each column and Latin hypercube an object having the same property in more than two dimensions. Consider the case of k variables. We want to use a number of samples equal to n . To perform a Latin hypercube sampling the cumulative distribution function of each variable of interest x_i is divided into segments of equal probability and a value is randomly selected in each interval as can be seen in fig. 1-9 where the cumulative distribution function CDF is shown together with the corresponding probability density function PDF.

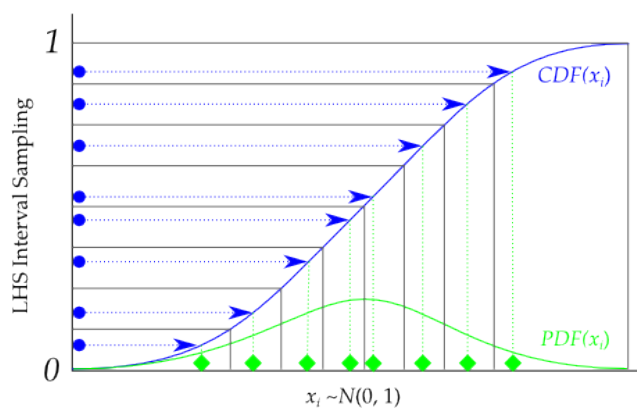


Figure 1-9: *Division of the cumulative probability of the variable x_i to perform a Latin hypercube sampling*

n samples are created assembling random combinations of the different values of x_i . An extensive description of this type of sampling, used for uncertainty and sensitivity analysis and compared to Monte Carlo technique can be found in [10].

Chapter 2

Surrogate Model

Once the inputs that have a significant impact on f are detected and the parameter space is investigated, a proper surrogate model should be chosen. Many types of surrogate models are available in literature.

An important class are the physically-based surrogates, built from a coarse model of the problem. This model can be constructed using a coarse discretization or simplifying the model equations for example neglecting some second-order terms. An advantage of the adoption of physically-based surrogates is that they rely on a-priori knowledge about the system of interest and can be subsequently improved if additional data from the high-fidelity model is available.

The second class of surrogate models can be call black-box approach: the high-order or full model problem is not modified and the surrogate is constructed using statistical techniques, interpolation methods or regression techniques. It is important to underline that this is not a rigid distinction but just a classification inside a relatively new research field. Among this second black-box approach we recall Kriging, response surface methods (RSMs) and Proper Orthogonal Decomposition (POD). The last two surrogate models are combined to construct the metamodel applied in

the current work and will be explained together.

2.1 Kriging

Kriging is a statistic interpolation technique born in geostatistics [15]. A Kriging model is described as

$$f(\mathbf{x}) = B(\mathbf{x}) + Z(\mathbf{x}), \quad (2.1)$$

where $B(\mathbf{x})$ is a trend function, decided a priori, and $Z(\mathbf{x})$ is a normally distributed Gaussian function with zero mean and variance σ^2 .

The covariance matrix of $Z(\mathbf{x})$ can be expressed as

$$COV[Z(\mathbf{x}_i), Z(\mathbf{x}_j)] = \sigma^2 R \quad (2.2)$$

and R is a $n \times n$ correlation matrix, with n equal to the number of samples. The form of the correlation matrix should be decided a priori and a common choice is to use the Gaussian correlation function proposed by Sacks [23]

$$R(x_i, x_j) = e^{\sum_{l=1}^n \theta_l |x_1 - x_2|^2}, \quad (2.3)$$

where θ_l are unknown correlation parameters. The correlation matrix R contains information about the weighted distance between the samples.

The trend function $B(\mathbf{x})$ in eq. 2.1 can be written as

$$B(\mathbf{x}) = g(\mathbf{x})^T \boldsymbol{\beta} \quad (2.4)$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_k]^T$ are unknown model parameters and $g(\mathbf{x}) = [g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_k(\mathbf{x})]$ contains low order polynomial, typically equal to constant functions.

Usually the outputs of Kriging interpolation are the kriging predictor $s(\mathbf{x})$, a mean expected value of the point of interest, equal to the average of all possible realizations, and a Kriging estimation of its variance. The Kriging predictor can be written as

$$s(\mathbf{x}) = g(\mathbf{x})^T \boldsymbol{\beta} + \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{f} - \mathbf{G}\boldsymbol{\beta}) \quad (2.5)$$

where $\mathbf{r} = [R(x, x_1), R(x, x_2), \dots, R(x, x_n)]^T$, \mathbf{f} is a vector containing the values of the function $f(\mathbf{x})$ in the samples, and \mathbf{G} is a $[n \times k]$ matrix with components $G_{ij} = g_j(\mathbf{x}_i)$.

The vector of model parameters $\boldsymbol{\beta}$ should be estimated as

$$\boldsymbol{\beta} = (\mathbf{G}^T \mathbf{R}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{R}^{-1} \mathbf{f}. \quad (2.6)$$

The variance of the Kriging predictor $\hat{\sigma}^2$ can be estimated as

$$\hat{\sigma}^2 = \frac{1}{k} (\mathbf{f} - \mathbf{G}\boldsymbol{\beta})^T \mathbf{R}^{-1} (\mathbf{f} - \mathbf{G}\boldsymbol{\beta}). \quad (2.7)$$

The model fitting is performed computing a maximum likelihood estimate for θ_k . Kriging method therefore is able to combine regression technique together with statistical concepts and it is able to foresee the areas in the parameter space where the errors of the model will be higher. It is used in plenty of disciplines as electrical and computer engineering, biology and telemetry.

2.2 Proper orthogonal decomposition

In this section the surrogate model used in the present work is described. The model relies on the Proper Orthogonal Decomposition (POD), a statistical technique able to extract the essential physics of some input information. The statistical analysis of the data allows to express the flow field in terms of a set of low rank basis vectors

and the surrogate model is obtained using this reduced set. In literature the POD appears in different equivalent forms such as Karhunen-Loève decomposition, Principal Component Analysis (PCA) and Singular Value Decomposition(SVD) [2]. According to POD an optimal linear basis, named POD basis, may be interpreted as a solution of a minimization of the projection error of the original system [13].

The POD has been applied in many different fields: data compression, image processing, dynamical systems, fluid mechanics. Its application in fluid dynamics was first introduced by Lumley [14] for the detection of coherent structures in a turbulent flow. In fluid dynamics, the POD is usually employed to find a basis for the projection of the Navier-Stokes equations and to obtain a physically-based surrogate model composed by a system of ordinary differential equations for the time dependent POD expansion coefficients [21]. Less commonly, the POD is applied in the frequency space [7] or in a parameter space. In this latter case, as an example, the POD can be used to describe flow fields around modified body shapes, using the information about the flow past few selected geometries of the body. Examples of this approach can be found in the works of Legresley and Alonso [12], Bui-Thanh et al. [6], Mifsud et al. [16] and Tang and Shyy [26].

The proper orthogonal decomposition technique can be seen as a mathematical procedure able to build a basis $\{\varphi_j(x)\}_{j=1}^{\infty}$ in order to construct a linear approximation $\tilde{f}(\mathbf{x})$ of the function $f(\mathbf{x})$:

$$\tilde{f}(\mathbf{x}) = \sum_{j=1}^M C_j \varphi_j(\mathbf{x}) \sim f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^k. \quad (2.8)$$

In this framework the basis $\{\varphi_j(x)\}_{j=1}^{\infty}$ is indicated as POD basis and the functions $\varphi_j(x)$ are called POD functions or modes. k is the cardinality of the parameter space and C_j are the POD coefficients. Many procedures are available in literature

to compute the POD basis. In this work we followed the method of snapshots [24]. A set of M representative samples $\{\mathbf{f}^i(\mathbf{x})\}_{i=1}^M$ named snapshots is considered.

The functions composing the basis $\{\varphi_j(x)\}_{j=1}^\infty$ are chosen to describe the M snapshots $\{\mathbf{f}^i(\mathbf{x})\}_{i=1}^M$ in an optimal way in the energetic sense, in fact they maximize the squared mean projection of the set $\{\mathbf{f}^i(\mathbf{x})\}_{i=1}^M$ on the basis itself. This can be represented by the optimization problem

$$\max_{\varphi} \frac{\langle |(\mathbf{f}(\mathbf{x}), \varphi(\mathbf{x}))|^2 \rangle}{\|\varphi(\mathbf{x})\|^2}, \quad (2.9)$$

that the POD basis has to satisfy. In Eq. (2.9) $\langle \cdot \rangle$ is the average operator, (\cdot, \cdot) indicates the inner product between two functions, $|\cdot|$ is the absolute value and $\|\cdot\|$ the L^2 norm.

For practical applications the technique should be used in finite dimensions. In this case the set of functions $\{\mathbf{f}^i\}_{i=1}^M$ becomes a set of M vectors and the basis $\{\varphi_j\}_{j=1}^M$ is no longer composed by functions but by vectors. The kernel necessary to solve the maximum problem (2.9) in finite dimensions is the correlation matrix R of dimensions $[M \times M]$. In this work the samples $\{\mathbf{f}^i\}_{i=1}^M$ consist of numerical solutions of a PDEs system and the superscript i indicates the varying boundary conditions [12]. To compute the basis vectors is necessary to solve the eigenvalue problem

$$R\mathbf{e} = \lambda\mathbf{e}, \quad (2.10)$$

where with R is indicated the correlation matrix

$$R_{ij} = \frac{1}{M}(\mathbf{f}^i, \mathbf{f}^j) \quad i = 1, \dots, M \quad j = 1, \dots, M. \quad (2.11)$$

Once the eigenvalue problem is solved, M eigenvalues $\{\lambda_i\}_{i=1}^M$ and M eigenvectors, each one with M components $\{\mathbf{e}_i^k\}_{i=1}^M, k = 1, \dots, M$ are obtained. The optimal POD

basis can be calculated recombining the initial snapshots with the eigenvectors and each basis vector will be constructed as

$$\varphi^j = \sum_{i=1}^M e_i^j \mathbf{f}^i \quad j = 1, \dots, M. \quad (2.12)$$

The physical meaning of the eigenvalues differs with the content of the vectors $\{\mathbf{f}^i\}_{i=1}^M$. Typically they represent a field variable. In the case of an incompressible flow for example, these vectors can consist of the discretization of the velocity field obtained from numerical solutions of the Navier-Stokes equations with different boundary conditions and each eigenvalue will be equal to the double of the mean kinetic energy captured by the corresponding POD mode [12].

The ratio

$$\frac{\sum_{i=1}^t \lambda_i}{\sum_{i=1}^M \lambda_i}, \quad (2.13)$$

where t corresponds to the number of modes considered, can be used as an indicator of the energy contained in the different modes. The general approach of the POD technique considers this eigenvalue estimate. Typically an a priori threshold is chosen and the modes corresponding to the eigenvalues with an amount of energy under the threshold are neglected. With this procedure, a truncation error is introduced and only q modes are considered, with $q \ll M$.

In the present work, a highly reduced number of snapshots (and therefore of POD modes) is used and no truncation is generated, avoiding the introduction of the corresponding error. All the POD modes are retained in the building of the POD surrogate model because the saving of computational cost performing the truncation was negligible.

Once the POD basis is constructed, it is possible to obtain exactly the initial snapshots $\{\mathbf{f}^i\}_{i=1}^M$ as the linear combination

$$\mathbf{f}^j = \sum_{i=1}^M C_i^j \boldsymbol{\varphi}^i \quad j = 1, \dots, M. \quad (2.14)$$

The coefficients C_i^j can be calculated as angles between the snapshots and the POD basis vectors:

$$C_i^j = (\mathbf{f}^j, \boldsymbol{\varphi}^i) \quad j = 1, \dots, M. \quad (2.15)$$

In fluid dynamics usually the truncated POD basis is used to project the Navier-Stokes equations along the optimal POD vectors and the mean operator is applied in time. In this way the reduced order model is formed by a set of ordinary differential equations in the unknowns $C_i^j(t)$.

In the present work a different approach is used: the POD is parametrized, no projection onto a reduced order dimension space is performed, the eigenvectors are all used to construct the POD basis and weighted with the coefficients C_i^j . In literature this particular application is called POD with interpolation or PODI [25].

With the PODI approach we follow the method of snapshots.

Let's call $\{\mathbf{f}^{\delta_i}\}_{i=1}^M$ the snapshot set corresponding to the parameter combination set $\{\boldsymbol{\delta}_i\}_{i=1}^M$. The parameter vector $\boldsymbol{\delta}$ has cardinality k and the M snapshots are related to M different parameter combinations. It is possible to obtain the POD basis $\{\boldsymbol{\varphi}_j\}_{j=1}^M$ following the method of snapshots described previously: a correlation matrix R should be calculated and an eigenvalue problem of order M has to be solved. The exact reconstruction of each snapshot is given by the linear combination

$$\mathbf{f}^{\delta_i} = \sum_{j=1}^M C_j^{\delta_i} \boldsymbol{\varphi}^j, \quad i = 1, \dots, M. \quad (2.16)$$

If we can assume that the coefficients $C_j^{\delta_i}$ are the discretization of the function $C_j(\boldsymbol{\delta})$ and if the function $C_j(\boldsymbol{\delta})$ is smooth enough with respect to $\boldsymbol{\delta}$, interpolation methods

can be used to calculate the PODI coefficients $\widetilde{C}_j^{\delta_l}$ for values of the parameters not belonging to the initial snapshot set.

Once the PODI coefficients $\widetilde{C}_j^{\delta_l}$ are computed, it is possible to evaluate the field variable $f_{un}^{\delta_l}$ corresponding to unknown parameter combinations δ_l as

$$f_{un}^{\delta_l} \sim \sum_{i=1}^M \widetilde{C}_i^{\delta_l} \varphi^i. \quad (2.17)$$

Eq. (2.17) is the representation of the surrogate model built in the present work and used for some of the following applications. The Navier-Stokes equations can be considered as the high order model of the problem, reference or 'true' solution. Starting from M numerical solutions of the Navier-Stokes equations, or snapshots, the surrogate model can be used to evaluate the field of a generic variable of interest (as density, temperature or the turbulent quantities) associated to specific boundary conditions corresponding to different parameter combinations δ_l not belonging to the initial snapshot set. No extrapolation procedure is allowed at the moment therefore δ_l has to stay in the parameter space defined by the snapshots. The use of the surrogate model instead of solving the system of PDEs allows to save computational time and effort maintaining acceptable accuracy with respect to the true solution. The accuracy is guaranteed by the fact that basis truncation is avoided and that the POD basis vectors are optimal in the energetic sense.

The method can be outlined in three main steps:

1. Generation of the initial snapshot set: the snapshots are obtained with high-fidelity CFD calculations. The position of the snapshots in the parameter space as well as their number strongly influence the ROM accuracy. This is the most computational expensive step but can be performed off-line only once.

2. POD decomposition: the correlation matrix R is computed and an eigenvalue problem of order M , equal to the snapshot number, is solved. As previously explained, in our approach M is small and all the POD modes are used to construct the surrogate model. There is no particular effort in the solution of the eigenvalue problem. Even this part can be performed one time off-line.
3. PODI reconstruction: this last step should be performed on-line for each reconstruction. The PODI coefficients $\tilde{C}_i(\delta_l)$ are computed using different interpolation techniques. For parameter spaces with dimension greater than one, the Response Surface Method is applied, with first and second order least square regression or Radial Basis Functions with gaussian or multiquadric bases. Once the coefficients are calculated, the linear combination of the POD basis can be performed in order to obtain the desired $f_{un}(\delta_l)$.

2.2.1 Response Surface Methodology

In the case of parameter spaces with dimension greater than one, the response surface methodology is applied in order to compute the PODI coefficients $\tilde{C}_i^{\delta_l}$ presented in Sec. 2.2. This methodology allows interpolation in more than one dimension through the generation of an analytical surface starting from a certain number of output evaluations.

The expression

$$\mathbf{y} = f(\mathbf{x}) + \epsilon \tag{2.18}$$

can be used [19] to approximate the input-output relation between \mathbf{x} and \mathbf{y} . In Eq. (2.18) \mathbf{y} is the dependent variable, called response of the system that we are analyzing and from that term arises the definition *response surface*. In our problem the vector

\mathbf{y} will contain the coefficients $\{C_j^{\delta_i}\}_{i=1}^M$ corresponding to the angles between the POD modes and the snapshots. For each a_j a different response surface has to be built, with $j = 1, \dots, M$; that is we are building M different response surfaces to approximate the PODI coefficients corresponding to each POD mode. The vector \mathbf{x} contains the independent variables or parameters of the problem. The error ϵ with respect to the true solution is modeled as a random error with 0 mean.

The function $f(\mathbf{x})$, used to approximate \mathbf{y} , can be composed by a certain combination of the components of the parameter vector \mathbf{x} . The complexity of this combination varies with the accuracy needed in the specific problem and in this work least square regression of the first and second order is used, together with radial basis functions considering gaussian and multiquadric basis.

Least Square regression

In the framework of response surface methodology, least square regression can be a simple technique to estimate a best fit approximation of the PODI coefficients $\tilde{C}_i^{\delta_l}$.

The general form of a first order least square surface can be written as:

$$y_i = \beta_0 + \sum_{j=1}^r \beta_j x_{ij} + \epsilon_i, \quad i = 1, \dots, M \quad (2.19)$$

The quantities indicated with β_j are called regressors and shall be estimate solving the system in the least square sense. The quantity r represents the number of regressors and therefore the number of unknowns of the system.

In the case of a second order regression we have

$$y_i = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=2}^k \beta_{ij} x_i x_j + \epsilon_i, \quad i = 1, \dots, M. \quad (2.20)$$

where k is the cardinality of the parameter space. The second order model can be useful in the case of a strong curvature in the true input-output relation.

In order to solve the problem, equation (2.19) or (2.20) can be written in matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (2.21)$$

with $\mathbf{y} \in \mathbb{R}^M$ and $\mathbf{X} \in \mathbb{R}^{M \times r}$. The vector $\boldsymbol{\beta}$ contains the r regressors of the model that are the unknowns of the system. M is the number of available input-output relations, or the number of already performed evaluations of the system state. In our case M is coincident with the number of snapshots.

System (2.21) can be solved through least square minimization

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \quad (2.22)$$

Once the vector $\boldsymbol{\beta}$ that contains the regressors of the model is found, the response surface is defined and the PODI reconstruction coefficients $\tilde{C}_i^{\delta_l}$ can be calculated directly from (2.19) or (2.20).

Radial Basis Functions

With the radial basis function method we can avoid regression and perform a classical interpolation, in order to build an analytical surface that is coincident with the data in the starting points, corresponding to evaluations of the high-order model. Radial basis functions can be considered a generalization of splines to the multivariate setting and they are able to avoid the curse of dimensionality treating all space dimensions in the same way.

The general form of a response surface built with radial basis functions is

$$y(x) = \sum_{j=1}^M w_j \phi(\|x - x_j\|_2), \quad x \in \mathbb{R}^k. \quad (2.23)$$

As we can see from Eq. (2.23), different positions of the radial function $\Phi = \phi(\|\cdot\|)$, with $\phi[0, \infty) \rightarrow \mathbb{R}$ produce the linear combination describing the response surface. The term radial is referring to the Euclidean norm $\|\cdot\|_2$. In our approach the vector $y(x)$ contains the known PODI coefficients $\tilde{C}_i^{\delta_l}$, M is the number of snapshots and w_j are unknown coefficients.

Starting from this general form many particularisations can be and have been made [5]. In this work we tested only two different basis forms: the gaussian basis $\phi(x) = e^{-\frac{1}{2\sigma_i^2}\|x-x_i\|_2^2}$ and the multiquadric basis $\phi(x) = \sqrt{1 + \frac{\|x-x_i\|_2^2}{2\sigma_i^2}}$. In these expressions σ is a shape parameter that determines the aspect of the radial basis function.

The solution of the interpolation problem (2.23) leads to the system

$$A\mathbf{w} = \mathbf{y}. \quad (2.24)$$

The interpolation matrix A has components $A_{ji} = \phi(\|x_i - x_j\|_2)$, $j, i = 1, \dots, M$. To avoid the problem of ill-conditioning a relaxation of the interpolation condition can be foreseen [4]. In this case data points and centers of the RBF functions are no longer coincident. The exact problem (2.23) becomes a problem of linear optimization that can be solved in the least square sense introducing the Moore-Penrose pseudoinverse. A more recent method to improve the conditioning of the interpolation problem is the addition of a multi-variate polynomial $p(x)$ leading to

$$y(x) = \sum_{j=1}^M w_j \phi(\|x - x_j\|_2) + p(x), \quad x \in \mathbb{R}^k. \quad (2.25)$$

Problem (2.25) is underdetermined and we have to impose the orthogonality condition [16]

$$\sum_{j=1}^M w_j p(x_j). \quad (2.26)$$

2.3 Proper Orthogonal Decomposition applied to the Parameter Space

In this section a novel approach in the construction of a surrogate model with proper orthogonal decomposition is presented. The basic idea is to apply POD to the parameter space in order to reduce the number of parameters to deal with and perform the optimization in a space with lower dimensions.

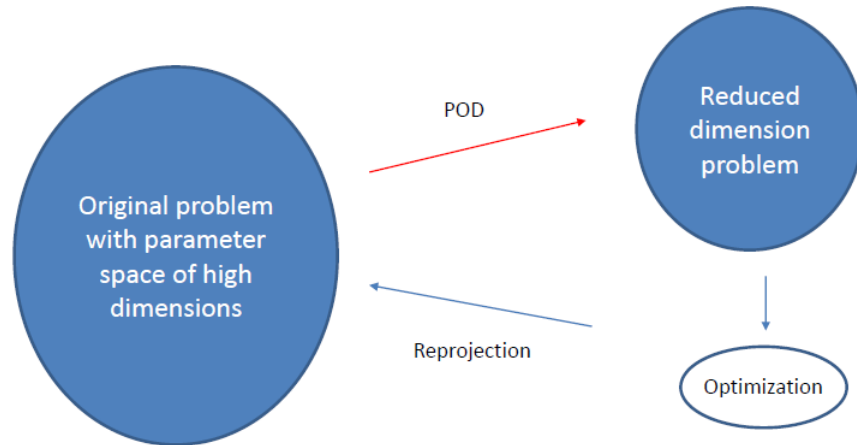


Figure 2-1: NACA0012 Basic idea of POD applied to the parameter space

2.3.1 Methodology

Let's consider the problem

$$p^* = \min_x f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^k \quad (2.27)$$

where p^* is the optimum value of the function $f(\mathbf{x})$ and k is the cardinality of the parameter space. Suppose that the parameter space is discretized with n points. Every point can be associated to a different value of the decision variable vector. It will exist therefore a set of vectors $[\mathbf{x}_1, \dots, \mathbf{x}_n]$.

To solve the minimization problem 2.27 several steps can be performed:

- creation of the $[n, k + 1]$ matrix V

$$V = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} & f_1(\mathbf{x}) \\ x_{21} & x_{22} & \dots & x_{2k} & f_2(\mathbf{x}) \\ \vdots & \vdots & \dots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} & f_n(\mathbf{x}) \end{bmatrix} \quad (2.28)$$

that describes the state of the system.

- apply POD to the matrix V :

In order to perform the decomposition, the correlation matrix between the columns of V should be constructed

$$R_{ij} = \frac{1}{k+1} (V_i, V_j), \quad i = 1, k+1, \quad j = 1, \dots, k+1. \quad (2.29)$$

From the POD decomposition of matrix R , $k+1$ eigenvectors \mathbf{e}_i , $i = 1, \dots, k+1$ and $k+1$ corresponding eigenvalues λ_i , $i = 1, \dots, k+1$ are obtained.

- computation of the reduction matrix \tilde{A}_r :

Consider the matrix

$$A_r = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{k+1}]. \quad (2.30)$$

The eigenvalues λ_i obtained from the POD decomposition can give a measure of the energy of the system that each eigenvector \mathbf{e}_i can contain. Therefore it can be possible to truncate the last q eigenvectors of the matrix A_r if they contain an energy value under an a-priori defined threshold and create the matrix \tilde{A}_r with dimensions $[k+1, k+1-q]$

$$\tilde{A}_r = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{k+1-q}]. \quad (2.31)$$

- projection of the matrix V in a reduced dimension space using matrix \tilde{A}_r :

$$\tilde{V} = V \cdot \tilde{A}_r. \quad (2.32)$$

The columns of the matrix \tilde{V} will contain the problem discretization in a reduced dimension space:

$$\tilde{V} = \begin{bmatrix} x_{11}^r & x_{12}^r & \dots & x_{1,k-q}^r & \tilde{f}_1(\mathbf{x}^r) \\ x_{21}^r & x_{22}^r & \dots & x_{2,k-q}^r & \tilde{f}_s(\mathbf{x}^r) \\ \vdots & \vdots & \dots & \vdots & \vdots \\ x_{n1}^r & x_{n2}^r & \dots & x_{n,k-q}^r & \tilde{f}_n(\mathbf{x}^r) \end{bmatrix}. \quad (2.33)$$

The optimization problem 2.27 in the reduced dimensions space will become

$$\tilde{p}^* = \min_{\mathbf{x}^r} \tilde{f}(\mathbf{x}^r), \quad \mathbf{x}^r \in \mathbb{R}^{k-q} \quad (2.34)$$

where the superscript r stays for reduced.

- identification of the minimum \tilde{p}^* in the reduced-dimension space:
a black-box optimizer can be used.
- re-projection of the optimum point in the original space:

$$\mathbf{m} = \mathbf{m}_r \cdot \tilde{A}_r^i \quad (2.35)$$

where \mathbf{m}_r is the $[1 \times k + 1 - q]$ vector containing the optimum point and the optimum value of the function $\tilde{f}(\mathbf{x}^r)$ in the reduced-dimension space, \mathbf{m} is the $[1 \times k + 1]$ vector containing the optimum point and the optimum value of the function in the original space and \tilde{A}_r^i is the pseudoinverse of matrix \tilde{A}_r .

Chapter 3

Applications

In this chapter applications of the two different surrogate model approaches previously exposed are presented. The first example is the simpler one and regards the backward facing step problem. A surrogate model using proper orthogonal decomposition for variable geometry is constructed and the influence of the number of snapshots is investigated. After that, surrogate models are built considering the problem of the flow around airfoil. Two typical shapes are considered: a NACA 00112 and a RAE 2822. In the first case therefore, a subsonic two-dimensional viscous flow is considered. Particular attention is given to the generation of the initial snapshot set. Different snapshot numbers and different distributions are tested. In the RAE 2822 problem on the other hand, we deal with a compressible problem and the shock wave analysis should be considered. The consecutive two applications are three-dimensional problems: the optimization of an automotive shape using a POD surrogate model and the construction of an aircraft surrogate model for database generation.

At the end of the chapter we find the minimization of an analytical function and contrarily from all the other cases here the POD is applied to the parameter space. Following this different approach the optimization is performed in a reduced-order

space and the optimum point is then re-projected in the original space.

3.1 Backward Facing Step

Part of the work described in this chapter is published in *Proper Orthogonal Decomposition as Surrogate Model for Aerodynamic Optimization*, V. Dolci and R. Arina, International Journal of Aerospace Engineering, Volume 2016 (2016), Article ID 8092824, 15 pages <http://dx.doi.org/10.1155/2016/8092824>

3.1.1 Problem setting

This test case is set according to the experimental work of Ruck and Makiola [22]. The

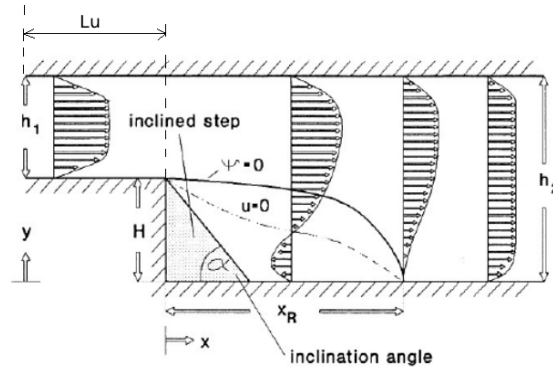


Figure 3-1: *Backward facing step with variable step slope*

flow enters in a channel from the inlet at a prescribed velocity and then encounters a step. The section of the channel increases causing the generation of a recirculation bubble in the flow. The length of the recirculation bubble is strongly dependent on length and slope of the step. A visualization of the problem setting is shown in Figure 3-1. The geometry of the problem is characterized by the Expansion Ratio (ER) that

is the ratio between the length of the channel section $(H+h_1)$ and the height of the inlet h_1 . In this work ER is equal to 2 with an inlet height $h_1 = 0.1$. In order to have a fully developed channel flow before the step, the value of the length L_u of the first part of the channel has to be chosen fulfilling the inequality $L_u > 5H$, therefore in this work L_u is equal to 1. Similarly, to have a fully developed flow in the channel behind the step, the length of the duct, measured from the step, is set to $40H$.

A set of four snapshots is obtained, modifying the slope angle α of the step. The angle α assumes the values of 90° , 45° , 30° and 25° . For each snapshot a full

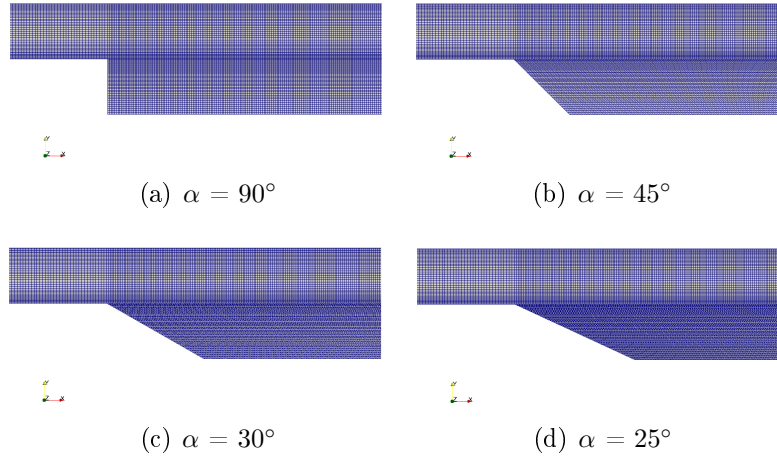


Figure 3-2: *Backward facing step. Geometries of the four snapshots*

CFD simulation is realized. The initial velocity is set equal to 2.5 m/s in all cases with a corresponding Reynolds number, referred to the height of the step H , $Re_H = 19800$. The simulation is done with the open source software OpenFOAM using the *simpleFoam* solver with a $k-\epsilon$ turbulence model. A visualization of the four different geometries used for the snapshot set is shown in Fig.3-2. The number of cells is about 350000.

The geometry variation is obtained through a morphing of the mesh from the base

configuration, keeping constant the number of cells.

3.1.2 POD reconstruction

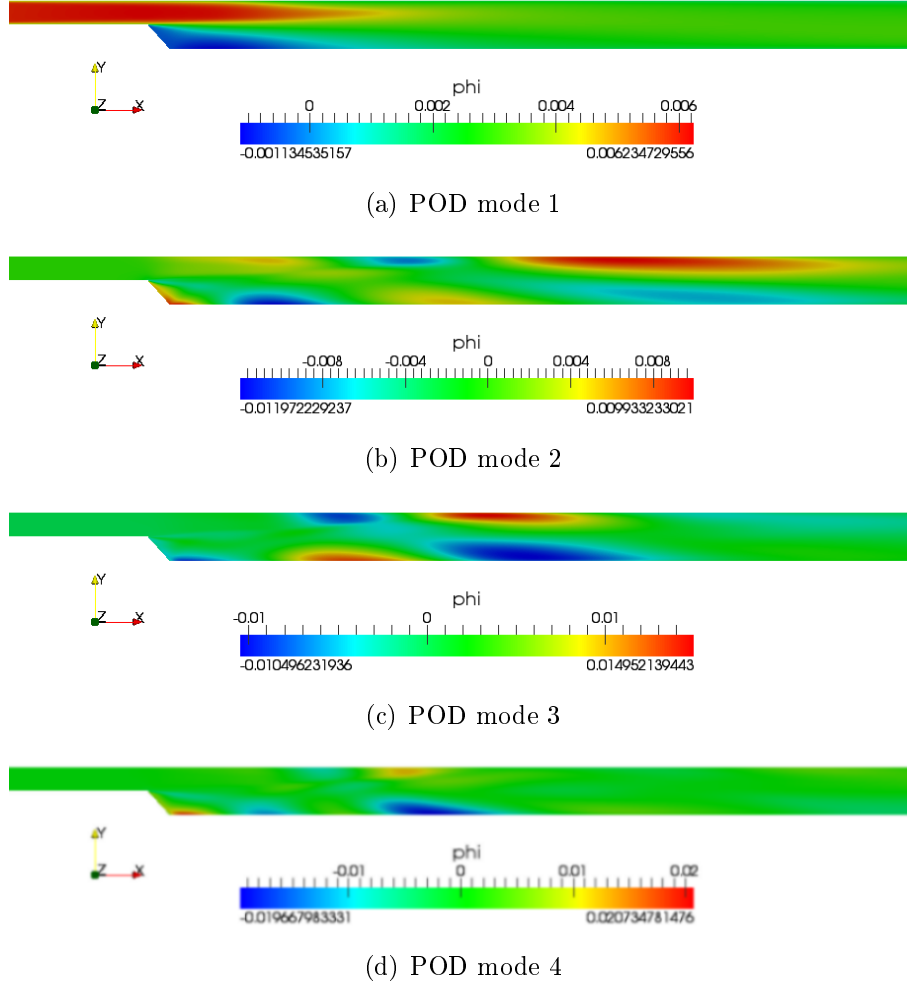


Figure 3-3: Backward facing step. POD modes for the velocity field

In Fig. 3-3 a visualization of the four POD basis vectors generated from the decomposition of the x-component of the velocity field is shown.

These basis vectors are the basic components of the surrogate model and are calculated as explained in section 2.2 Then velocity and pressure fields for a geometry characterized by a slope angle of the step of 48° , not belonging to the initial set of snapshots, are reconstructed using the POD surrogate model. The parameter space is one dimensional and linear interpolation is used to compute the POD coefficients. The POD reconstruction of the x-component of the velocity is shown in Fig.3-4, and compared with the field obtained with a CFD calculation. The L_2 percentage error,

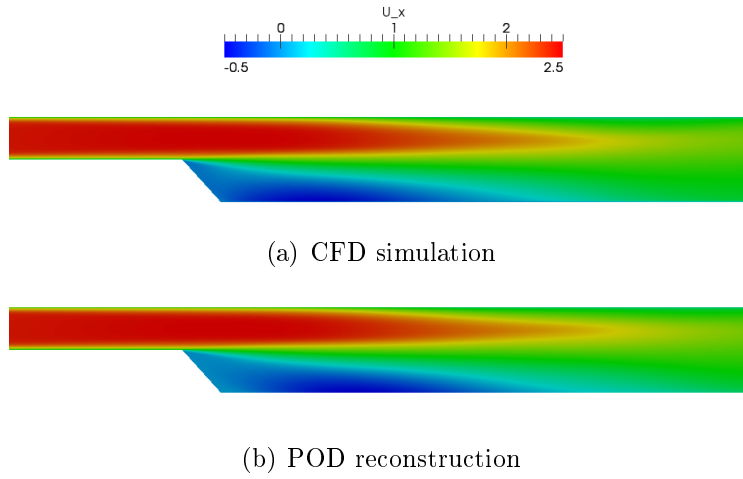


Figure 3-4: Backward facing step. x-component of the velocity [m/s], step slope $\alpha = 48^\circ$

expressed by

$$E\% = \|\mathbf{x} - \mathbf{x}_{POD}\|_2 \cdot 100 , \quad (3.1)$$

where with \mathbf{x} is indicated the value of the field of interest calculated with the CFD full model and with x_{POD} the corresponding value obtained with PODI, is $5.7e^{-4} \%$ for the pressure field and $8.2e^{-3} \%$ for the velocity field. The surrogate model therefore is able to predict the behavior of the system for the point $\alpha = 48^\circ$ in the parameter space, within an acceptable accuracy.

Analyzing Fig.3-3 and Fig.3-4, it can be seen that the first POD mode is qualitatively identical to the reconstructed field. An estimate based on Eq.(2.13) indicates that the first POD mode contains the 99.8 % of the total field energy. The fact that the first POD mode is able to represent the most field energy is intrinsic in the POD technique because it maximizes the mean of the norm of the squared field projection along the POD modes. Nevertheless in this work all the POD modes were used for the reconstruction because the saving of computational effort and time of retain fewer modes was negligible.

In addition, selecting two sections, one placed in the fully developed channel flow and the other in the recirculation bubble (Fig.3-5), it is possible to compare (Fig.3-

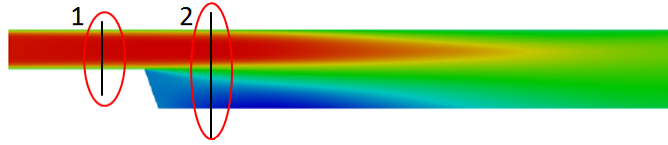


Figure 3-5: Backward facing step. Control sections

6) the velocity profile obtained with the POD reconstruction with the high-fidelity solution.

The maximum relative error $e\%$, calculated as

$$e\% = \max \left| \frac{x_{POD} - x}{x} \cdot 100 \right| , \quad (3.2)$$

is 1.18 % in the recirculation bubble. Probably the maximum error takes place in this area because is a low energy zone and the POD technique is optimal in the energetic sense.

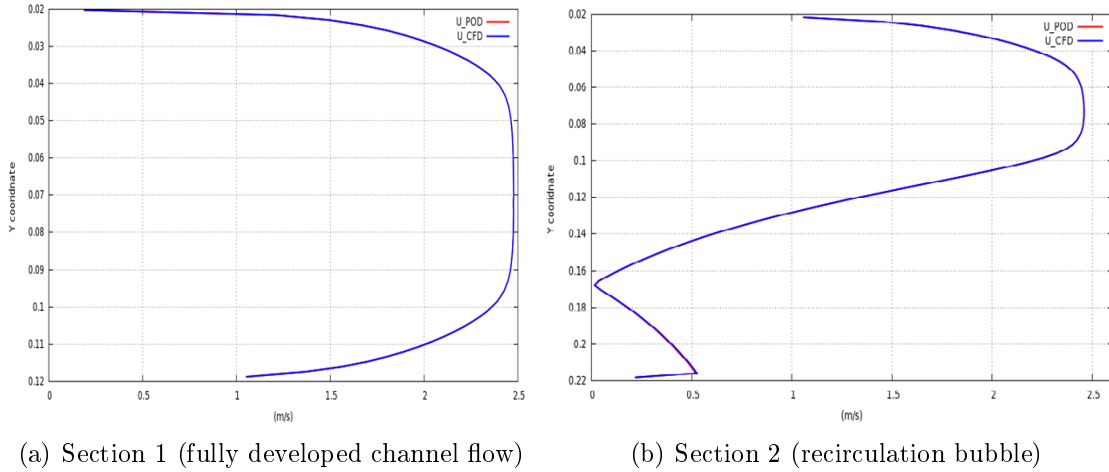


Figure 3-6: Backward facing step. Velocity magnitude profiles, $\alpha = 48^\circ$

3.1.3 POD reconstruction with different snapshot sets

To quantify the influence of the number of snapshots on the accuracy of the surrogate model, a further test is made. The problem setting is the same of section 3.1.1 but now four different surrogate models are constructed and compared using 3, 6, 11 and 21 snapshots. In this case the geometry is fixed for all the snapshots, with step slope of 90° , and each snapshot is calculated imposing a different inlet velocity. The inlet velocity varies in the range between 10 and 30 m/s. In table 3.1 a summary of the four sets of snapshots used in the different reconstructions is reported.

The surrogate model is used to reconstruct pressure and velocity fields for an initial velocity of 15.5 m/s not belonging to any of the initial sets of snapshots. Based on estimate (2.13), it can be remarked how the energy captured by the first mode increases with the number of snapshots composing the set. Moreover the first mode is always able to get more than the 99.99 % of the total energy.

A comparison between the reconstruction error generated by the four surrogate models for the pressure field is shown in Fig. 3-7. The error $|p - p_{POD}|$ is calculated

3 snapshots case		6 snapshots case	
Snapshot number	Initial velocity m/s	Snapshot number	Initial velocity m/s
1	10	1	10
2	20	2	14
3	30	3	18
		4	22
		5	26
		6	30
11 snapshots case		21 snapshots case	
Snapshot number	Initial velocity m/s	Snapshot number	Initial velocity m/s
1	10	1	10
2	12	2	11
3	14	3	12
4	16	4	13
5	18	5	14
6	20	6	15
7	22	7	16
8	24	8	17
9	26	9	18
10	28	10	19
11	30	11	20
		12	21
		13	22
		14	23
		15	24
		16	25
		17	26
		18	27
		19	28
		20	29
		21	30

Table 3.1: Backward facing step. Summary table of the snapshot parameters

along the center line of the duct. As expected this difference is decreasing as the number of snapshots increases. However in the three-snapshot case the maximum error is already acceptable and below $2.5 Pa$. In Fig. 3-8 the variation of the reconstruction error $\|e\|_{\infty} = \max(|p - p_{POD}|)$ of the pressure field using the four-snapshot set can be seen. The error trend is clear and the use of 11 snapshots for the surrogate model seems to be optimal, in terms of computational time and accuracy, compared with the 21-snapshot case. As further confirmation, in Fig. 3-9 a comparison between the reconstruction of the pressure field using 3 and 21 snapshots is reported. As expected, the result with many snapshots is more precise, but in the three-snapshot

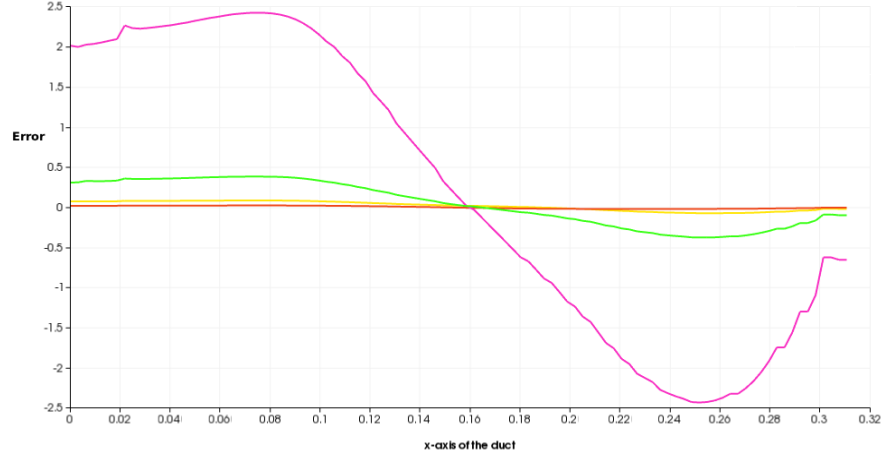


Figure 3-7: Backward facing step. $|p - p_{POD}|$ along the duct center line, for 3 (violet line), 6 (green line), 11 (yellow line) and 21 (red line) snapshots

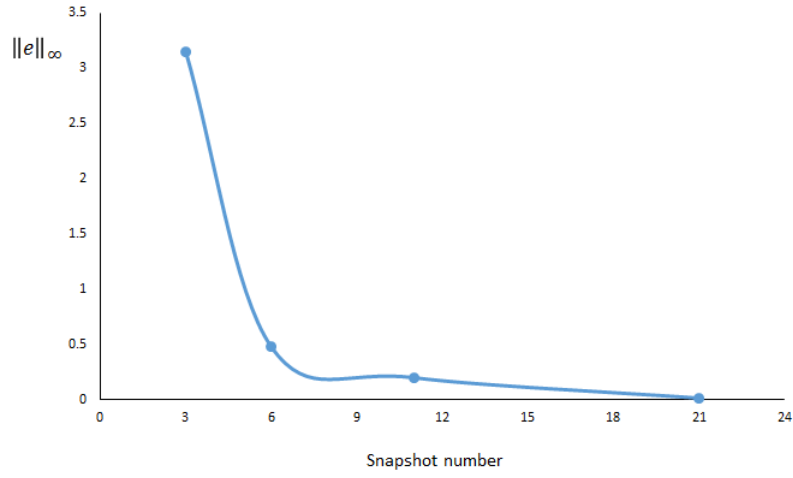


Figure 3-8: Backward facing step. Error $\|e\|_\infty$ on the pressure field with respect to the number of snapshots

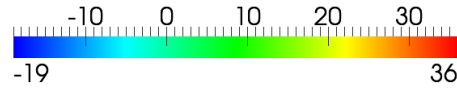
reconstruction the error is already acceptable. The POD decomposition is optimal in the energetic sense therefore in a steady problem the first POD mode is quite always able to get a relevant part of the field energy and a reduced number of modes and snapshots can be used to build a surrogate model.



(a) CFD solution



(b) POD reconstruction – 21 snapshots



(c) POD reconstruction – 3 snapshots

Figure 3-9: Backward facing step. Pressure field [Pa]

3.2 NACA 0012 Surrogate Model

Part of the work described in this chapter is published in *Variational Analysis and Aerospace Engineering - Mathematical Challenges for the Aerospace of the Future*, Frediani A., Mohammadi B, Pironneau O., Cipolla V., Springer Optimization and its Applications 2016.

A surrogate model is constructed considering the flow around a NACA 0012 airfoil. The model is used to evaluate unknown pressure and velocity fields corresponding to desired parameter configurations in the two-dimensional parameter space composed by the Mach angle of the undisturbed flow M and the angle of attack α of the airfoil. As previously explained the high-order model of the problem are the Navier Stokes equations. A steady subsonic flow is considered and a grid characterized by a number of points N equal to 260000 is used to discretize the geometry. To compute the snapshots the SIMPLE algorithm implemented in the open source software OpenFOAM is used. To model the turbulence a Spalart-Allmaras equation is solved and wall functions are used to estimate the boundary layer quantities near the wall.

In Fig. 3-10 a representation of the problem discretization can be seen and in Fig. 3-11 a zoom near the airfoil is shown.

3.2.1 Interpolation technique - preliminary phase

A first section of the construction of the surrogate model is dedicated to the choice of the right interpolation technique for the computation of the PODI coefficients $\tilde{C}_i^{\delta_l}$. For this purpose a snapshot set $\{\vec{f}^i\}_{i=1}^M$ composed of only 4 snapshots is used. The snapshots are placed at the four vertices of the parameter space and this configuration corresponds to a 2-level full factorial plane.

In Fig. 3-12 a representation of the snapshot position in the parameter space can

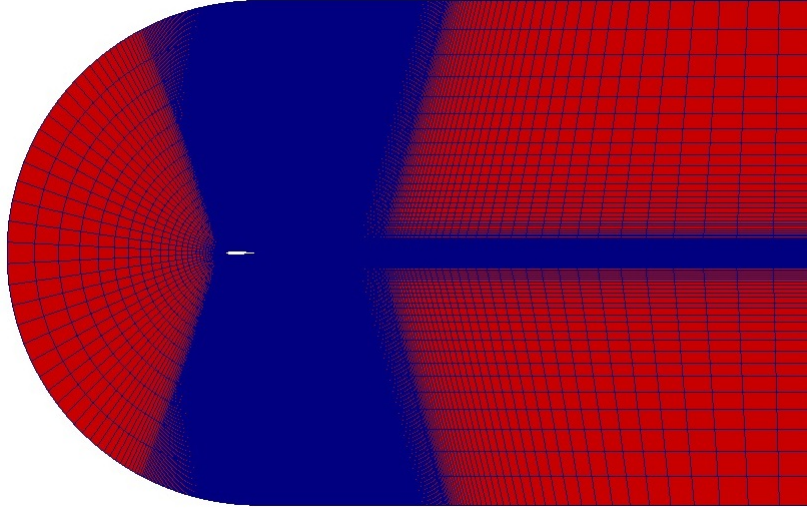


Figure 3-10: Problem discretization for the NACA 0012 surrogate model

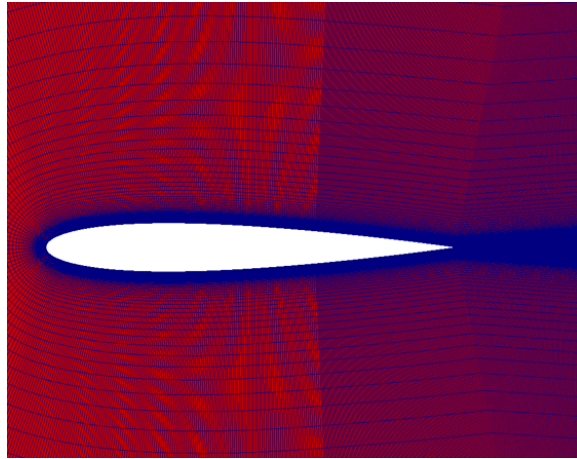


Figure 3-11: Problem discretization for the NACA 0012 surrogate model - airfoil zoom

be seen. Three points are randomly chosen to test the surrogate model and decide which interpolation technique has to be preferred for the model. The NACA 0012 surrogate model is only a preliminary application and this number can be increased

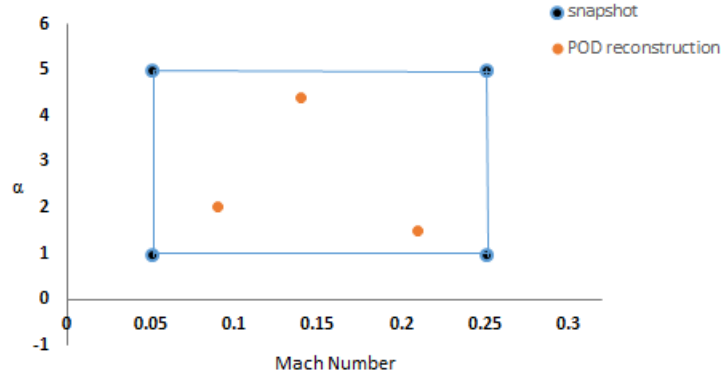


Figure 3-12: Snapshot and PODI reconstruction positions in the parameter space

in practical cases. It is necessary to consider that higher is the reconstruction number in this first section, higher will be the total computational cost to build the surrogate model. A trade off between accuracy and computational resources, in terms of hardware and time, shall be taken into account. If an a priori knowledge of the problem is available, it is possible to start testing the model in portions of the parameter space that will be more interesting or critical from the design point of view. In Fig. 3-12 the three test points can be seen in red.

Table 3.2: Snapshot and POD reconstruction positions in the parameter space

Snapshot	Mach Number	Alpha [°]
1	0.05	1
2	0.25	1
3	0.05	5
4	0.25	5
PODI reconstruction	Mach Number	Alpha [°]
1	0.09	2
2	0.21	1.5
3	0.14	4.4

In Table 3.2 snapshots and test points are reported. Four techniques are tested

to compute the PODI reconstruction coefficients $\tilde{C}_i^{\delta_l}$: a least square regression of the first order (LS), radial basis functions using gaussian basis with and without polynomial term and using multiquadric basis considering the polynomial term. A comparison of the results can be seen in Fig. 3-13. In this figure, only the error on

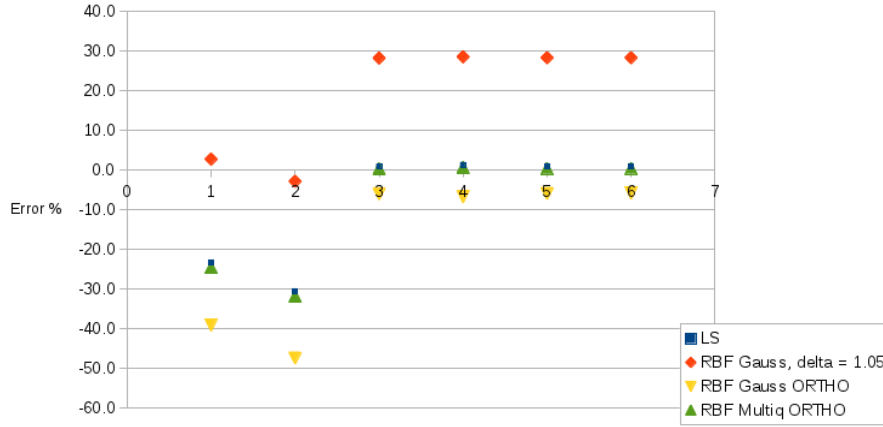


Figure 3-13: NACA0012 surrogate model. Errors for different response surfaces (x = 1: error on the maximum value of the pressure field, x = 2: error on the minimum value of the pressure field, x = 3 and 4: errors on maximum and minimum values of the x component of the velocity field, x = 5 and 6: errors on maximum and minimum values of the y component of the velocity field)

the maximum and minimum values of pressure and velocity fields associated to the three reconstruction points 1, 2 and 3 are considered.

The error $e_{\%}$ is calculated with respect to the CFD computation, taken as reference or 'true' solution and is computed as

$$e_{\%} = \frac{x - x_{POD}}{x} \cdot 100 . \quad (3.3)$$

The best results for the reconstruction of the pressure field are obtained using a radial basis function with gaussian basis, no relaxation of the interpolation condition

and a value of the shape parameter σ of 1.05. In this application a constant parameter σ is chosen for all the different radial basis functions therefore $\sigma_i \equiv \sigma, i = 1, \dots, M$. The behaviour of the response surfaces is different in the PODI reconstruction of the velocity field. In this case the lower error is obtained using a multiquadric basis, with the orthogonality condition. These interpolation methods will be used for the airfoil surrogate models built in the following section to reconstruct pressure and velocity fields.

3.2.2 Influence of the snapshot number

Once the best interpolation technique has been selected for the PODI coefficients, the influence of the number of snapshots on the accuracy of the surrogate model can be investigated. Four PODI surrogate models are tested using 4, 9, 16 and 25 snapshots corresponding to a 2, 3, 4 and 5 level full factorial design for the two parameters Mach and α .

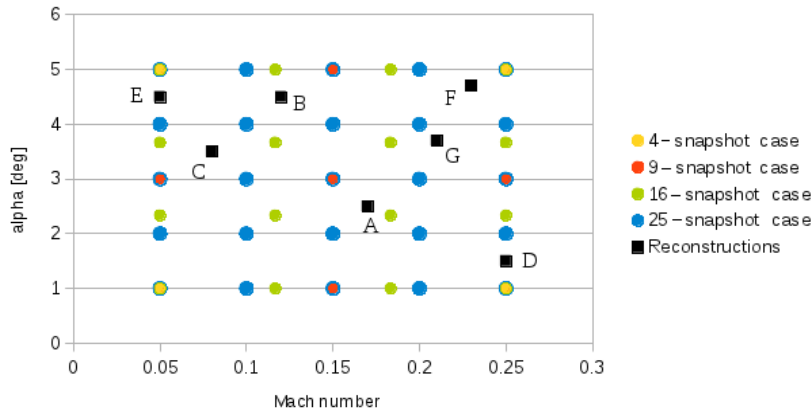


Figure 3-14: NACA0012 surrogate model. Snapshot sets and reconstruction positions in the (Mach- α) plane

With respect to the surrogate model of the previous section, the parameter ranges are extended: the Mach number now is varying from 0.05 to 0.25 and α is between 1° and 5° . In Fig. 3-14 a visualization of the snapshot positions is shown. The surrogate models are used to reconstruct pressure and velocity fields in seven random points in the parameter space. The positions of the reconstructed cases in the (Mach- α) plane are summarized in Table 3.3.

Table 3.3: NACA0012 surrogate model. List of the reconstruction combinations

Reconstruction point	Mach number	α [$^\circ$]
A	0.17	2.5
B	0.12	4.5
C	0.08	3.5
D	0.25	1.5
E	0.05	4.5
F	0.23	4.7
G	0.21	3.7

In this case the error $E\%$ with respect to the CFD reference solution is computed using the L_2 norm

$$E\% = \|x - x_{POD}\|_2 \cdot 100 = \frac{\sqrt{\sum_{i=1}^N (x_i - x_i^{POD})^2}}{N} \cdot 100, \quad (3.4)$$

where with x is indicated the CFD value of the field of interest in a single cell and with x^{POD} the corresponding value obtained with PODI.

In Figs. 3-15 and 3-16, the error trends obtained with Eq. (3.4) are shown for the pressure and velocity fields respectively. As expected the errors are decreasing with the increasing of the snapshot number. Considering the velocity field, the error is already under 2 % using 4 snapshots and is slowly decreasing. Therefore, depending on the a priori threshold of the surrogate model error, the use of 25 snapshots can be avoided and 9 or 16-snapshot sets can be used. For the pressure field higher errors are

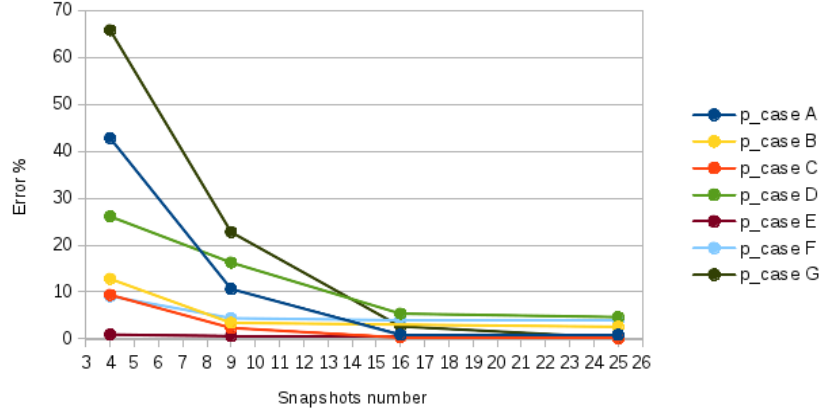


Figure 3-15: NACA0012 surrogate model. Errors generated by the PODI-surrogate model for the reconstruction of the pressure field

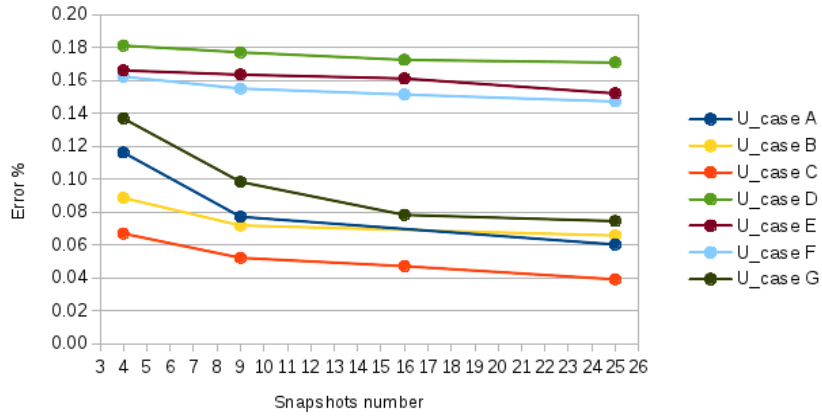


Figure 3-16: NACA0012 surrogate model. Errors generated by the PODI-surrogate model for the reconstruction of the velocity field

generated, but again the use of the 16-snapshot set can fulfill accuracy requirements. In Fig. 3-17 a visualization of the velocity field is shown for the full model and for the 25, 16 and 9-snapshot cases for point C.

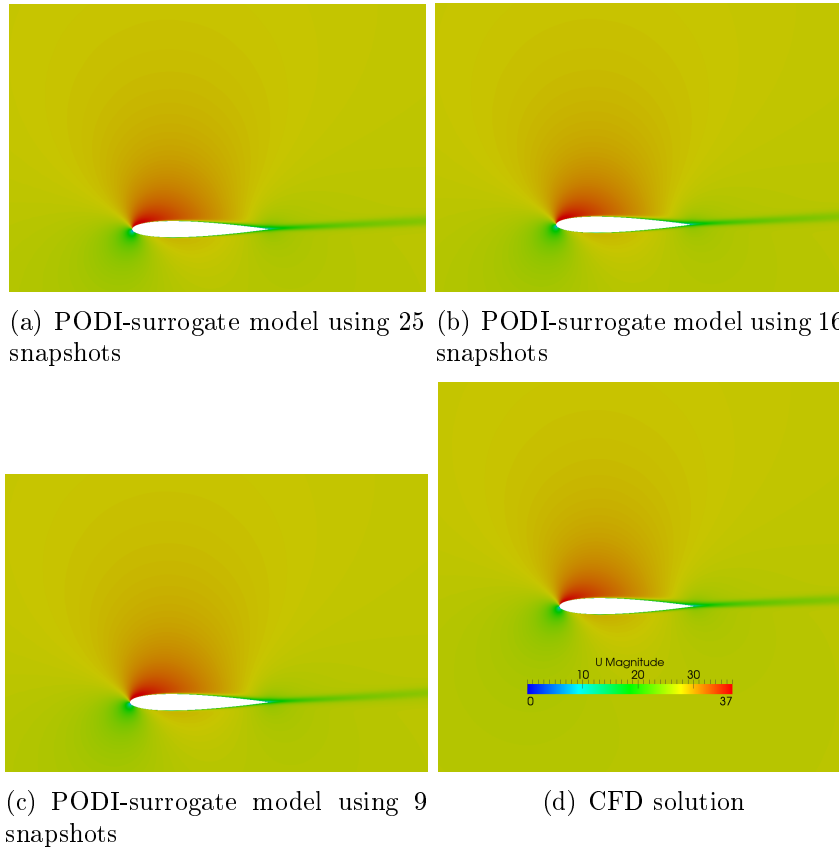


Figure 3-17: NACA0012 surrogate model. Comparison between velocity fields. $M = 0.08$, $\alpha = 3.5^\circ$

3.2.3 Influence of the snapshot position - Quadtree initial distribution

As previously explained, Quadtree is a specific subdivision of 2-dimensional spaces used first in digital imaging and examples of a quadtree distribution applied to the initial snapshot sampling for the building of a surrogate model can be found in [8]. In the present work no leave-one-out procedure is present because we are using a reduced number of snapshots.

In the Figs. 3-18 and 3-19 a comparison between 5-level full factorial and quadtree distribution can be made and in Figs. 3-20 the results are reported.

The quadtree distribution is able to reduce the surrogate model errors in all the seven test points. This can be obvious for the internal points B, C, A and F, since in the quadtree distribution the snapshots are now nearer to the reconstruction point but is not trivial for the other points D, E and G. In this three points, with respect to the 5-level case, the snapshots in the quadtree distribution are farther but still the errors are decreasing.

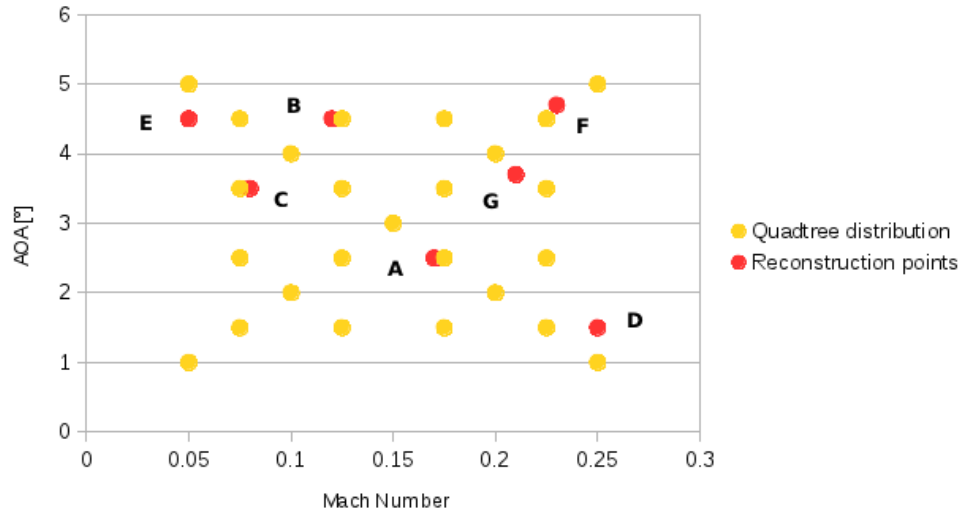


Figure 3-18: NACA0012 surrogate model. Snapshot and reconstruction positions in the 5-level full factorial case

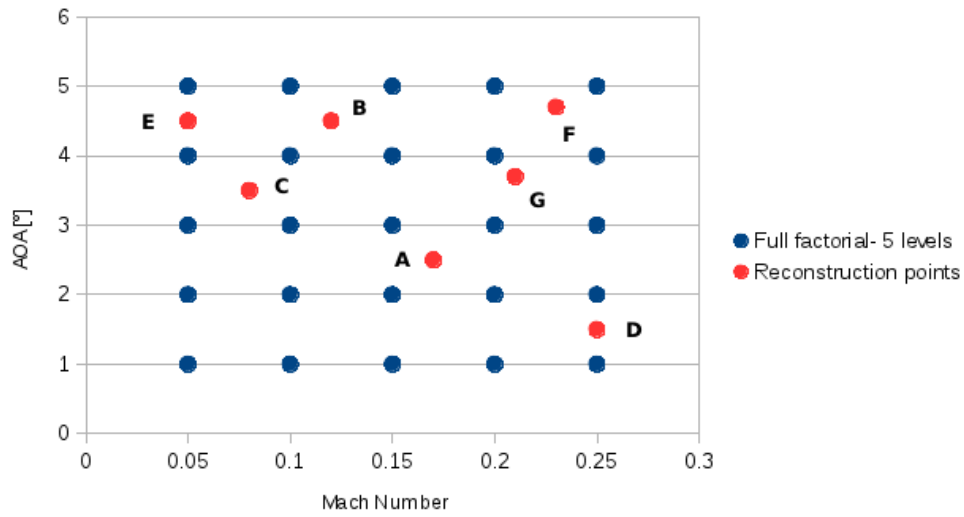


Figure 3-19: NACA0012 surrogate model. Snapshot and reconstruction positions in the quadtree case

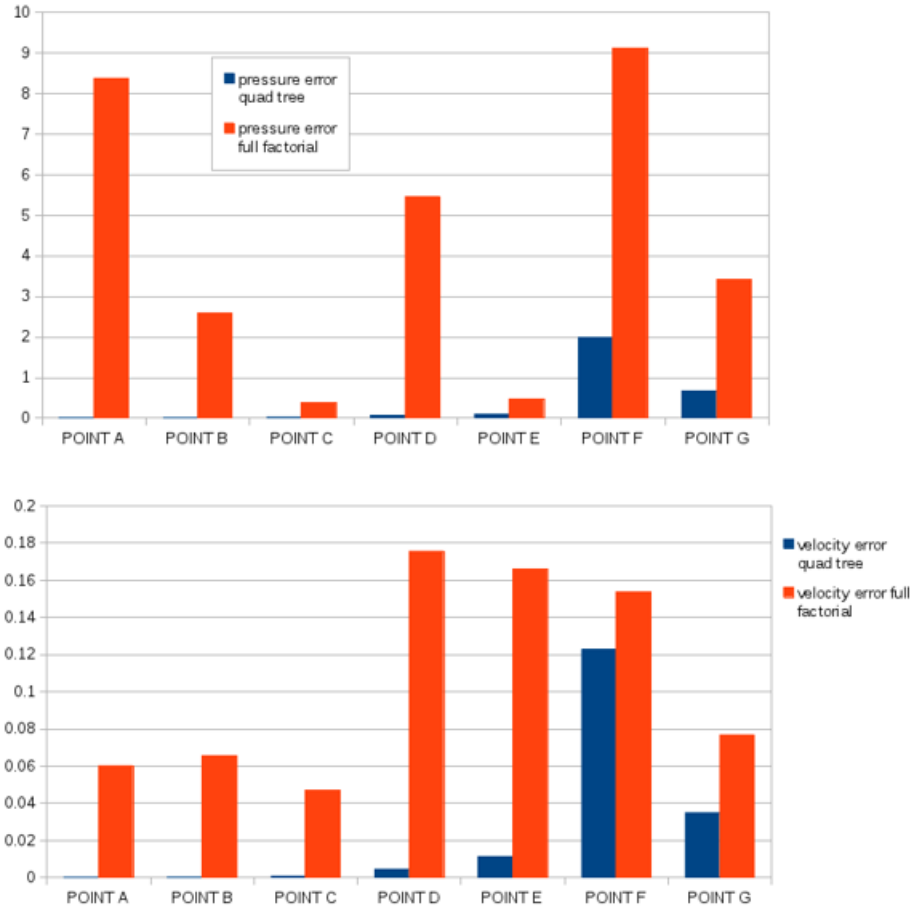
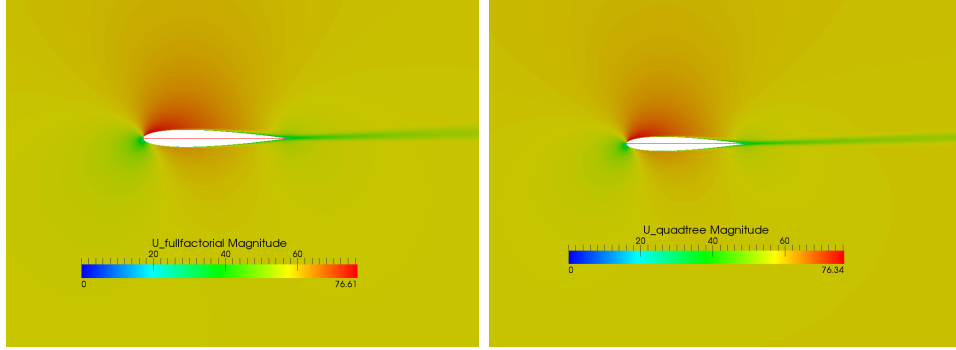
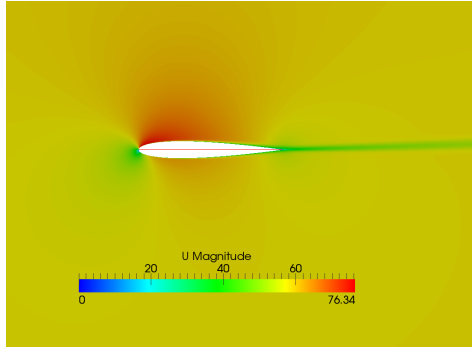


Figure 3-20: NACA0012 surrogate model. Comparison of the surrogate model errors on the field reconstruction using quadtree or 5-level full factorial distribution. Pressure (top) and velocity (bottom) fields

Finally, in Fig. 3-21, visualizations of the reconstructed fields compared with the CFD high order solution can be seen.



(a) Velocity field obtained with the reduced-order model using a 5-level full factorial distribution (b) Velocity field obtained with the reduced-order model using a quadtree distribution



(c) Velocity field obtained with the CFD full model

Figure 3-21: Velocity fields, m/s

3.3 RAE 2822 Surrogate Model

In this section the generation of a surrogate model for a supercritical airfoil is performed. The flow past a transonic wing profile is complex and highly nonlinear and particular care must be taken in the shock wave treatment.

3.3.1 Problem Setting

The two-dimensional transonic flow field past the supercritical airfoil RAE 2822 is analyzed. A surrogate model is built in the shape parameter space. The airfoil is modeled using two Bézier curves [9], one for the upper surface and one for the lower surface of the airfoil. Each Bézier curve is computed using 9 control points as can be seen in Fig. 3-22.

The high-order simulations are performed considering a flow field characterized by a Mach number of 0.729 and a Reynolds number of $6.5 \cdot 10^6$. The airfoil has an angle of attack equal to 2.31° . The geometry was discretized using 2500 cells. A steady flow was considered, with a $k-\omega$ EARSM (Explicit Algebraic Reynolds Stress Model) [20] model to describe the turbulence.

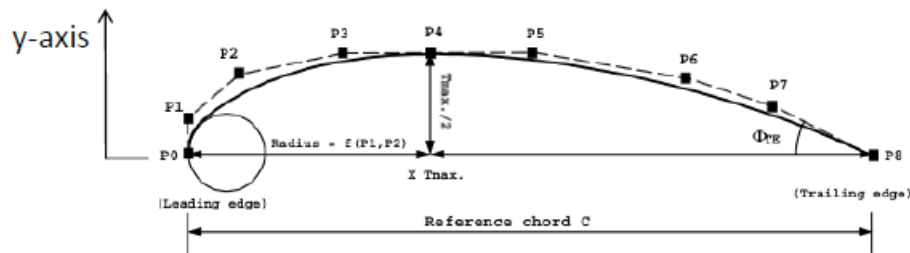


Figure 3-22: RAE 2822 surrogate model. Bézier curve for the upper surface of the airfoil. Visualization of the 9 control points

To construct the parameter space of the surrogate model, only 5 control points are considered. The points P0, P1, P7 and P8 are fixed to maintain a constant position of leading and trailing edge. Points P2, P3, P4, P5, P6 instead are moved along the y-axis, normal to the flow field. The parameters of the surrogate model are the positive or negative variations of the point positions along the y-axis with respect to the base configuration, within a range of ± 2 of the chord. The parameter space is 5-dimensional and we are affected by the curse of dimensionality: if a 2-level full factorial plane is required for a preliminary screening of the space, 2^5 simulations should be performed. In order to reduce the number of initial simulations required for the construction of the surrogate model, a 2-level fractional factorial plane 2^{5-1} is adopted to compute the snapshots. In this way only 16 high-order model evaluations are required instead of 32. Fractional factorial planes are a useful instrument typically applied in DOE (design of experiment) methodology. They allow to investigate the response with respect to multiple parameters with a reduced number of samples [19]. In Table 3.4 the 16 parameter combinations associated to each snapshot are reported.

3.3.2 Results

In this application the POD surrogate model is used to reconstruct the conservation variables (density ρ , momentum $\rho\mathbf{u}$, and total energy ρE) and the coordinates of the grid points. A mesh morphing is used to build only the initial snapshots, for the remaining reconstruction points the computational mesh is computed with the PODI model. The reconstructed fields are associated to shape parameter combinations not belonging to the initial set of snapshots. 4 random points, listed in Table 3.5, are chosen to test the surrogate model. The reference solution is obtained through CFD simulation using the Alenia solver UNS3D. With respect to this solution, the normalized root mean square error E_n is computed for each reconstructed field as

Table 3.4: RAE 2822 surrogate model. List of the parameter combinations associated to the snapshots

Snapshot number	$\Delta P2$	$\Delta P3$	$\Delta P4$	$\Delta P5$	$\Delta P6$
1	-0.2	-0.2	-0.2	-0.2	-0.2
2	0.2	-0.2	-0.2	-0.2	-0.2
3	-0.2	0.2	-0.2	-0.2	-0.2
4	0.2	0.2	-0.2	-0.2	0.2
5	-0.2	-0.2	0.2	-0.2	-0.2
6	0.2	-0.2	0.2	-0.2	0.2
7	-0.2	0.2	0.2	-0.2	0.2
8	0.2	0.2	0.2	-0.2	-0.2
9	-0.2	-0.2	-0.2	0.2	-0.2
10	0.2	-0.2	-0.2	0.2	0.2
11	-0.2	0.2	-0.2	0.2	0.2
12	0.2	0.2	-0.2	0.2	-0.2
13	0.2	0.2	-0.2	-0.2	-0.2
14	0.2	-0.2	0.2	0.2	-0.2
15	-0.2	0.2	0.2	0.2	-0.2
16	0.2	0.2	0.2	0.2	0.2

Table 3.5: RAE 2822 surrogate model. List of the 4 reconstruction points

Point number	$\Delta P2$	$\Delta P3$	$\Delta P4$	$\Delta P5$	$\Delta P6$
1	0.2	0	0	0	0
2	0.1	0.1	0.1	0.1	0.1
3	0.2	0.1	0.2	0.1	0.2
4	0.15	-0.1	0.15	-0.1	0.15

$$E_n = \frac{\sqrt{n \sum_{i=1}^n (x_i - x_i^{POD})^2}}{\sum_{i=1}^n x_i} \cdot 100, \quad (3.5)$$

where n is the number of grid points, x the CFD value of a cell and x^{POD} the corresponding value obtained with the surrogate model. In Table 3.6 this error is

listed for the 4 points.

Table 3.6: RAE 2822 surrogate model. Errors on the reconstruction of the 4 test points

	Point 1	Point 2	Point 3	Point 4
Field variable	E_n %	E_n %	E_n %	E_n %
ρ	0.4	1.2	0.5	0.6
ρu	1.0	3.6	1.3	1.1
ρv	5.4	23.9	8.8	6.9
ρE	0.5	1.2	0.5	0.5

The values are high considering the transverse momentum especially for point 2. In Figs. 3-23, 3-24, 3-25, 3-26, qualitative comparisons are reported for derived fields of Mach and pressure.

A partial agreement with the reference solution is achieved. A relative error e_{rel} is calculated considering the computation of lift and drag coefficients c_L and c_D :

$$e_{rel} = \frac{|x_{CFD} - x_{POD}|}{x_{CFD}} \cdot 100 \quad (3.6)$$

and the results are listed in Table 3.7. If we take into account only drag and lift coefficient reconstructions, considering for example a link of the surrogate model to an optimization procedure, the errors e_{rel} are lower than the root mean square computation E_n and a good agreement with the CFD result is obtained.

Table 3.7: RAE 2822 surrogate model. Errors on lift and drag coefficients

	$e_{rel} (c_L)$	$e_{rel} (c_D)$
Point 1	0.3	0.1
Point 2	0.2	1.5
Point 3	0.4	0.4
Point 4	1.0	2.6

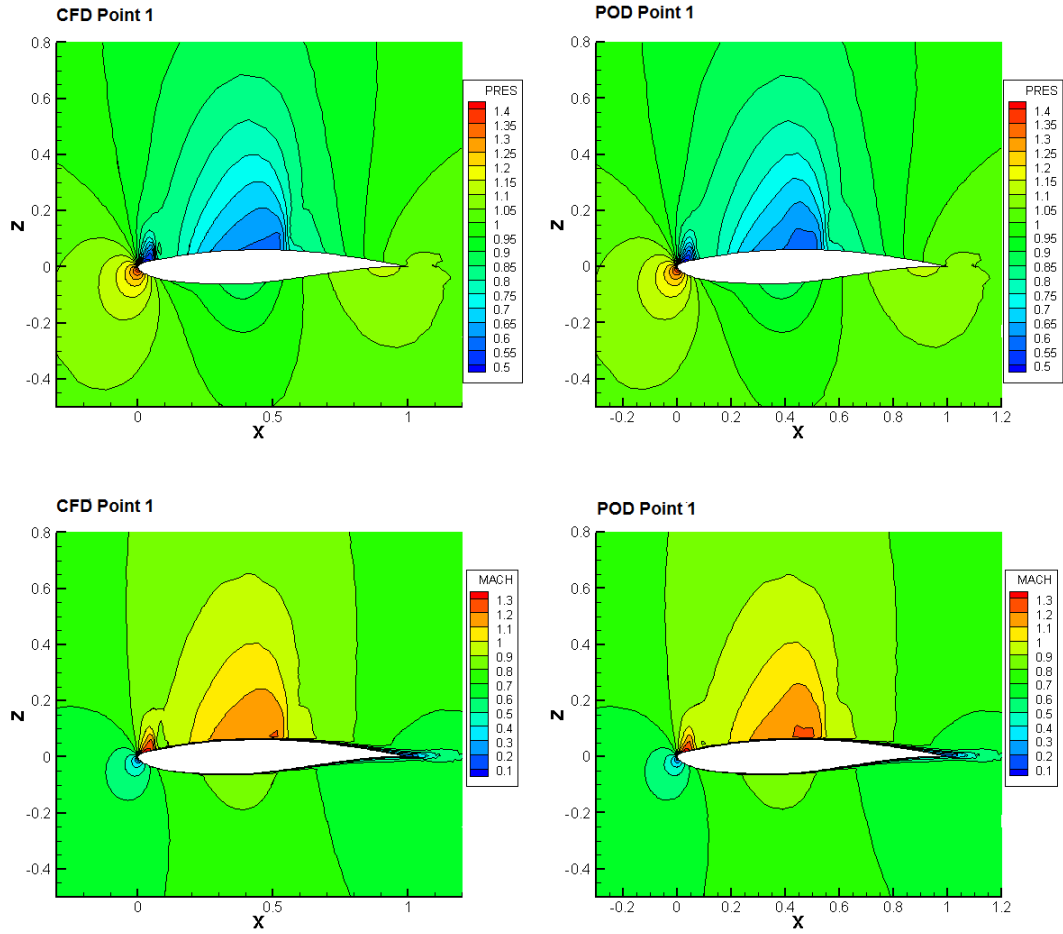


Figure 3-23: RAE 2822 surrogate model. CFD and PODI comparison for reconstruction point 1: pressure (top) and Mach (bottom) fields

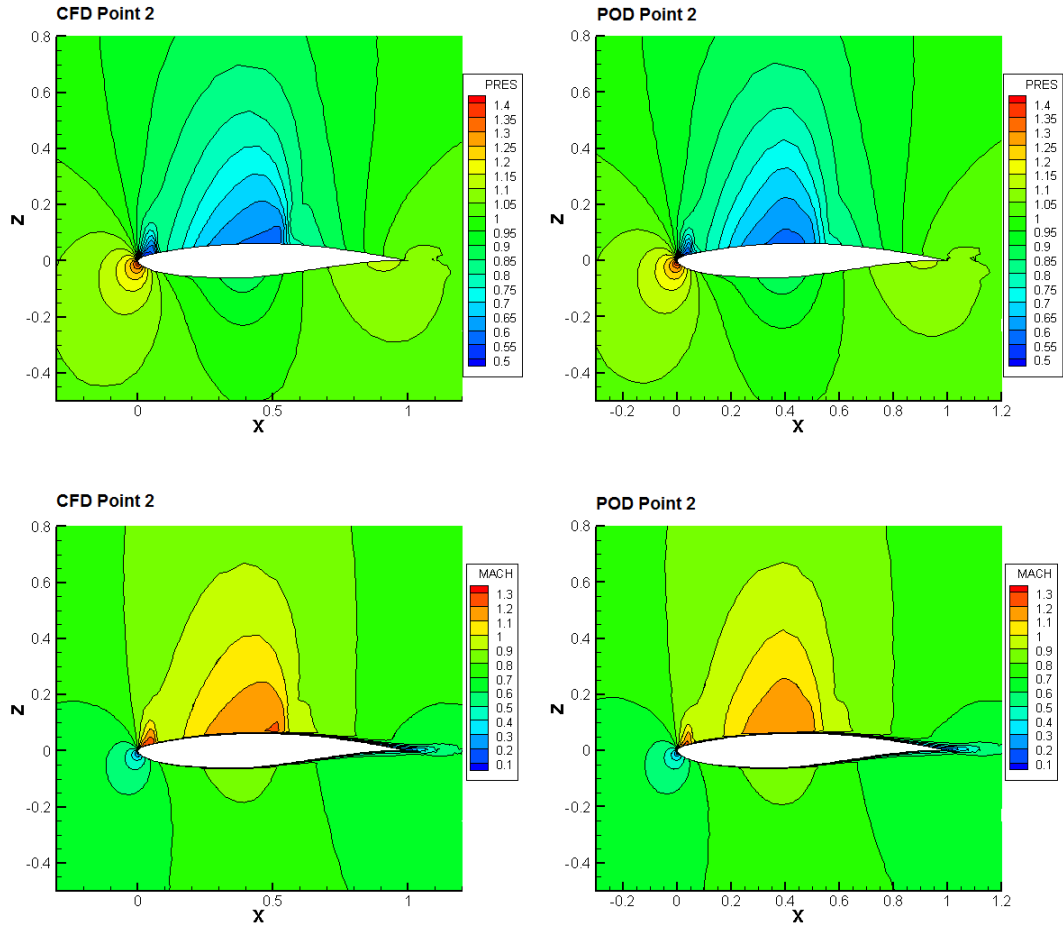


Figure 3-24: RAE 2822 surrogate model. CFD and PODI comparison for reconstruction point 2: pressure (top) and Mach (bottom) fields

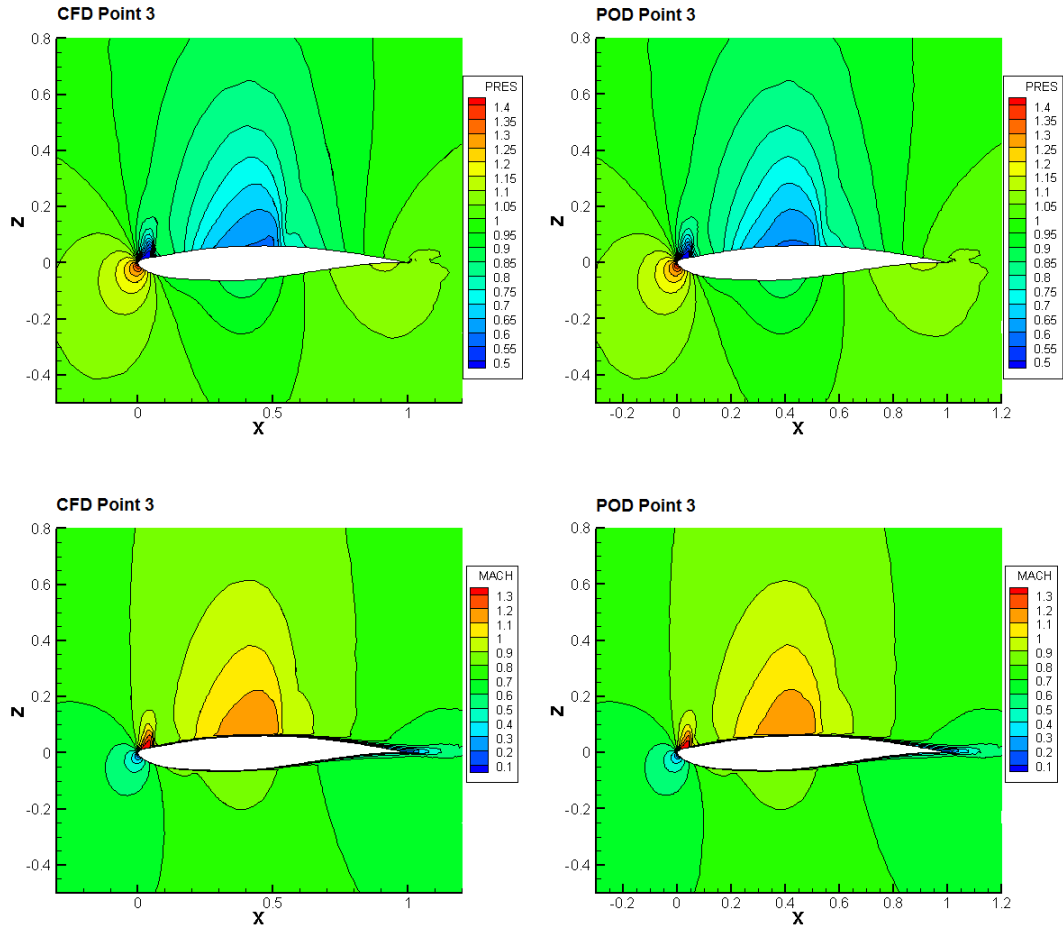


Figure 3-25: RAE 2822 surrogate model. CFD and PODI comparison for reconstruction point 3: pressure (top) and Mach (bottom) fields

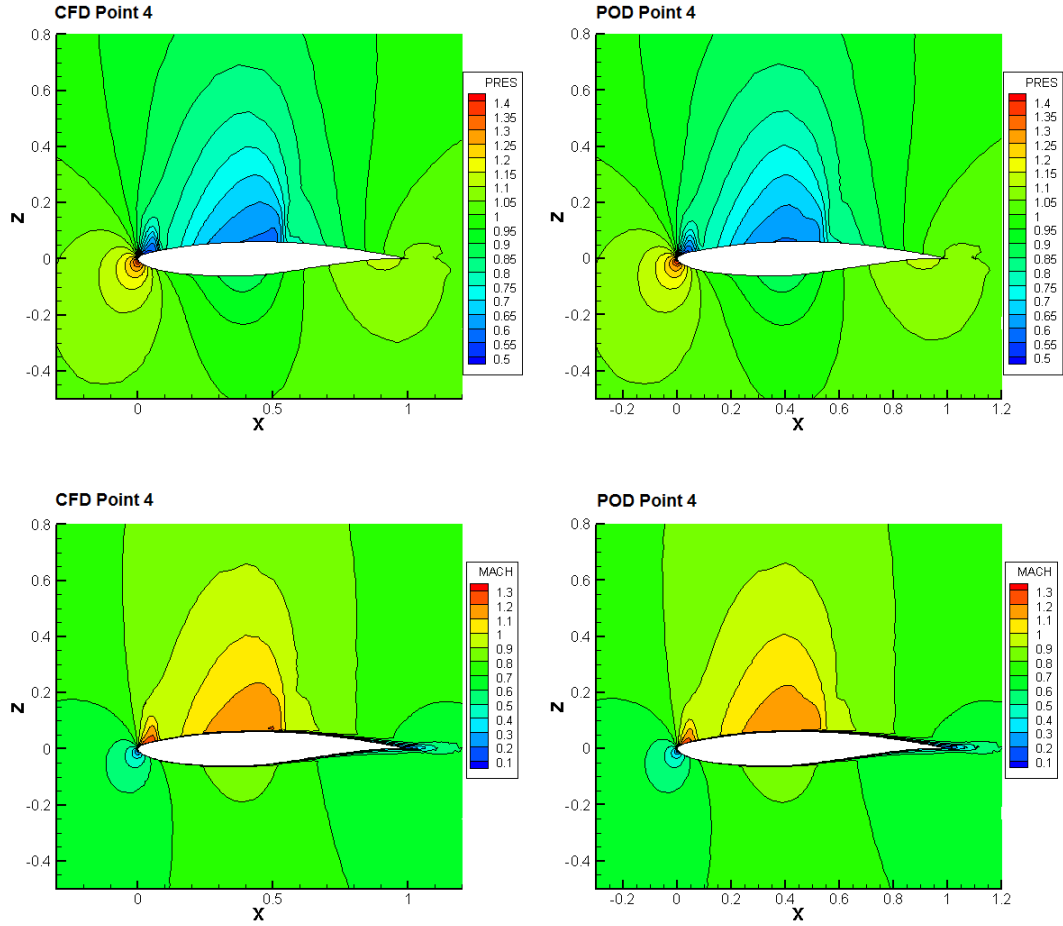


Figure 3-26: RAE 2822 surrogate model. CFD and PODI comparison for reconstruction point 4: pressure (top) and Mach (bottom) fields

3.4 Drag Coefficient Optimization of an Automotive Shape

3.4.1 Problem and CFD setting

The POD surrogate model is employed in an optimization loop to obtain the minimum drag coefficient of an automotive shape. The base form is the open source DrivAer car model from the Technical University of Munich [18]. The drag coefficient c_L is minimized acting on four shape parameters: the length x_1 , width y_1 and height z_1 of the trunk and the height of the diffuser z_2 (Fig. 3-27).

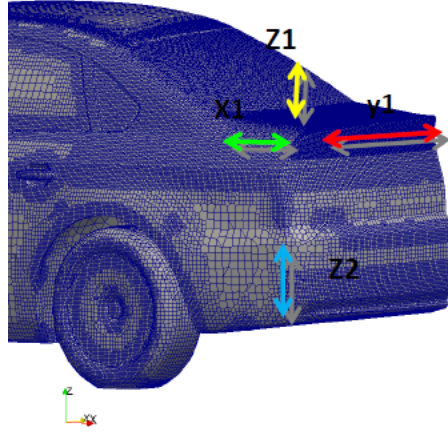


Figure 3-27: DrivAer model. Shape parameters chosen as design parameters

The optimization problem can be formulated as

$$\begin{aligned}
 c_L^* = \min_{x_1, y_1, z_1, z_2} & \quad \frac{D(x_1, y_1, z_1, z_2)}{\frac{1}{2} \rho_\infty U_\infty^2 A_{ref}} \\
 s.t. \quad & -0.5 \leq x_1 \leq 0.5, \\
 & -0.5 \leq y_1 \leq 0.5, \\
 & -0.4 \leq z_1 \leq 0.4, \\
 & -0.5 \leq z_2 \leq 0.5,
 \end{aligned}$$

with D being the vehicle drag, ρ_∞ the density and U_∞ the velocity of the undisturbed flow and A_{ref} a reference area equal to the vehicle maximum frontal area.

The geometry is discretized with 1959410 cells and the CFD simulation is performed with the OpenFOAM solver *simpleFoam* using a $k - \epsilon$ turbulence model. U_∞ is 40 m/s and the Reynolds number $Re_\infty \approx 14 \cdot 10^6$. The correct deformations corresponding to the desired values of the design variables are imposed with mesh morphing. Therefore the cell number remains constant for each geometry and only some point positions are modified.

3.4.2 Optimization

In this specific problem the cost function evaluations correspond to the calculations of the drag coefficient. Pressure, velocity, turbulence kinetic energy, specific turbulence dissipation, turbulence eddy viscosity ν_t and the mass flow through the cell faces φ are reconstructed. The drag coefficient is computed in a post-process phase. The selection of the set of snapshots is made on the basis of the Central Composite Design theory. The number of snapshots determines the total number of CFD calculations.

An appropriate selection of the snapshots is extremely important because their calculations represent the most time consuming step of the entire optimization procedure. The calculation of the POD basis is done following the method described in Section 2.2. Response surfaces are generated with a least square method of the second order or interpolating using radial basis functions as explained in Section 2.2.1.

The optimization has to be performed in a 4-dimensional parameter space therefore, willing to adopt a full level factorial design, we have to construct 2^4 snapshots only for the 2-level case and this number would grow exponentially increasing the levels. Taking into account the error trends of the surrogate models built in the previous test cases, a reduced initial snapshot set can be used in this practical case. A 2-level fractional factorial design 2^{4-1} , adding the central point, is adopted instead of a full level and the snapshot set is composed of only 9 snapshots. In Table 3.8 a description of the design parameter combinations used to generate the snapshots is reported.

Table 3.8: DrivAer model. 2-level fractional factorial design for the initial snapshot set

Snapshot number	x_1	y_1	z_1	z_2
1	−1	−1	−1	−1
2	−1	+1	+1	−1
3	+1	+1	−1	−1
4	+1	−1	+1	−1
5	−1	+1	−1	−1
6	−1	−1	+1	+1
7	+1	−1	−1	+1
8	+1	+1	+1	+1
9	0	0	0	0

In this table the highest and lowest values of the variables are represented with +1 and −1 respectively. The real values of the high and low levels of the design

parameters are determined by the optimization constraints and are ± 0.5 for x_1 , y_1 and z_1 and ± 0.4 for z_2 .

Once the snapshots are calculated, the remaining function evaluations required by the optimization algorithm are obtained using the PODI surrogate model.

The optimization algorithm used in this application is the SOGA (Single Objective Genetic Algorithm) implemented in the JEGA library of the open source software Dakota. SOGA is a classical single-objective genetic algorithm that performs optimizations of a single cost function. Obviously the use of a genetic algorithm is not mandatory: the PODI surrogate model can be linked to any other optimizer.

In this problem the design variables are represented in floating-point, with a random initialization and control to avoid duplications. The number of individuals composing the initial population is 50. A shuffle random crossover type is set with a rate of 0.8. With this particular crossover the parent chromosome sequences, once selected, are randomly shuffled and then the single-point crossover is performed. This operation is useful to eliminate the positional bias associated with the length of each chromosome. A random mutation rate of 0.08 is imposed. This kind of mutation corresponds to a random selection of an individual and a random selection of a design variable at which is assigned a random valid value. After 7 generations, the genetic algorithm is able to identify an optimal solution calculated using the surrogate model. The optimal drag coefficient is 0.3013, starting from a base configuration (all parameters at 0 level) of 0.3111, with an improvement of the 1.8 %. The error with respect to the CFD solution is 1.36 % for the optimal point.

The adoption of a PODI-surrogate model dramatically reduces the computing time of the optimization: instead of ~ 240 h without the adoption of the surrogate model, ~ 90 h are necessary with PODI, using 2 Intel Xeon E5440 Quad core processors.

Fig. 3-28 leads to considerations on the energy associated to the POD modes. The results have a direct analogy with the previous test cases and in particular with the

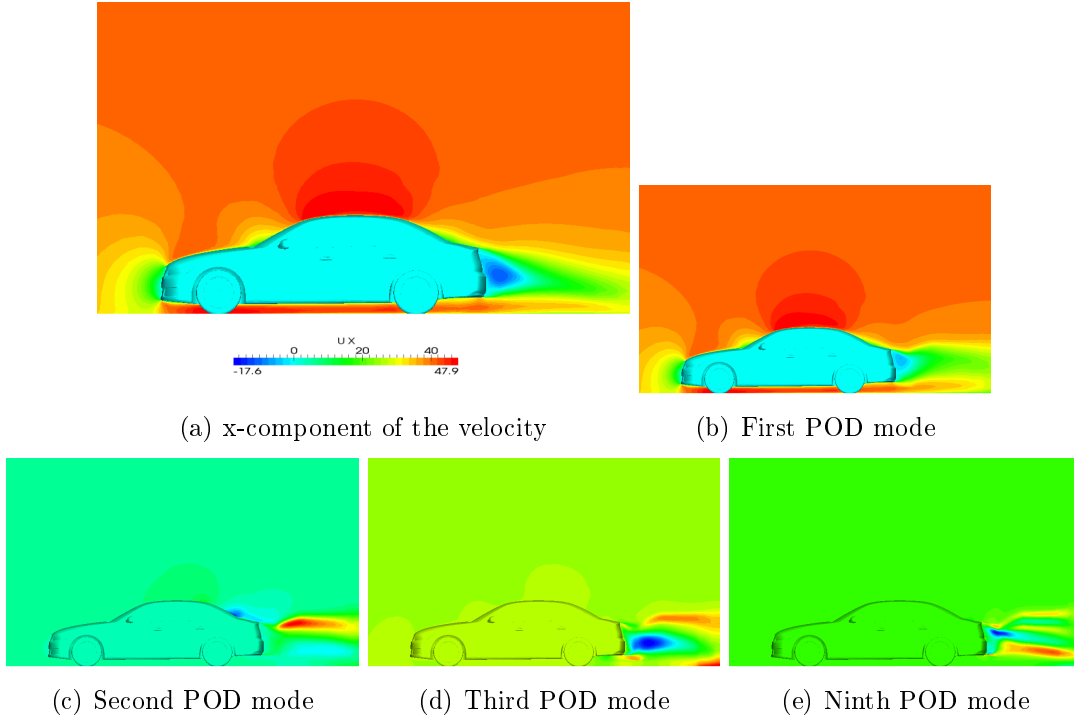
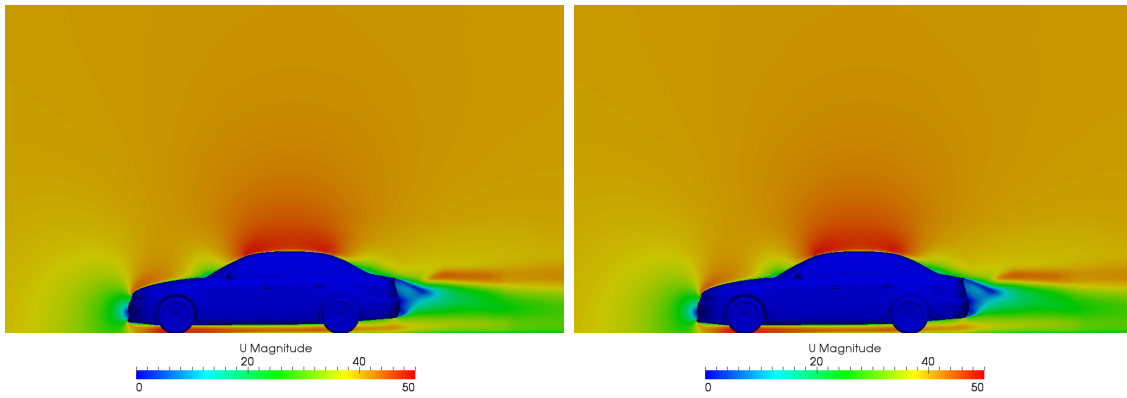


Figure 3-28: DrivAer model. Comparison between the velocity field obtained from the PODI-surrogate model and the individual POD modes

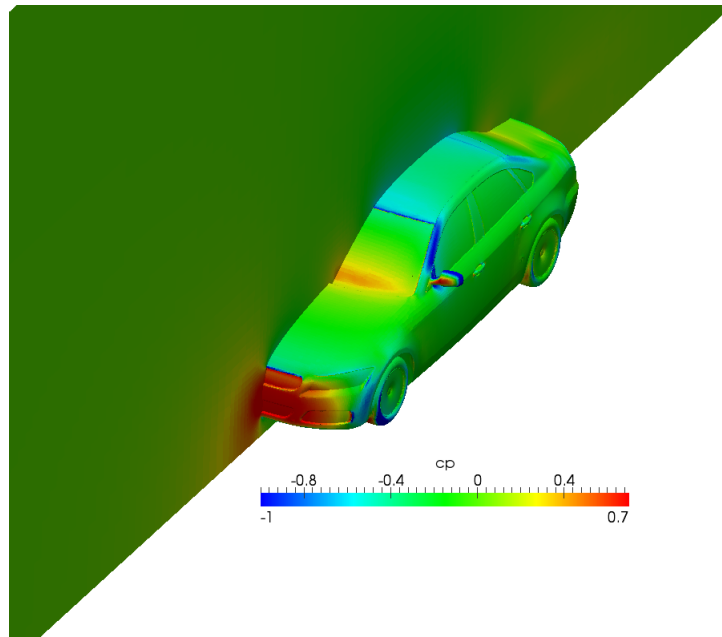
backward facing step problem. Considering for example the x-component of the velocity, the first POD mode is able to represent the greater amount (99.3 %) of the energy and its qualitative appearance is comparable to the reconstructed field. The other POD modes add information on the zones of the field with minimal energy, therefore mostly on the vehicle wake. All the POD modes are used for the surrogate model evaluation, as shown in the previous test cases, since this addition is not computationally demanding. In Figures 3-29 and 3-30 the comparisons between the velocity and pressure coefficient fields obtained with CFD solver and PODI reconstruction are shown.



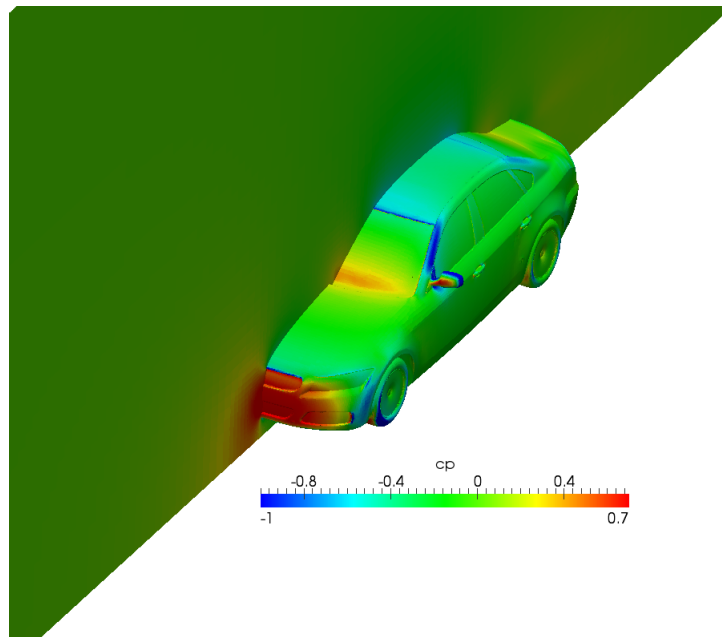
(a) CFD solution

(b) POD reconstruction

Figure 3-29: DrivAer model. Velocity fields (magnitude m/s)



(a) CFD solution



(b) POD reconstruction

Figure 3-30: DrivAer model. Pressure coefficient

3.5 3D Aircraft Surrogate Model

A PODI surrogate model is built for a three-dimensional aircraft. The two-dimensional parameter space is composed by the angle of attack α and the sideslip angle β of the aircraft. In Fig. 3-31 a geometry visualization is reported.

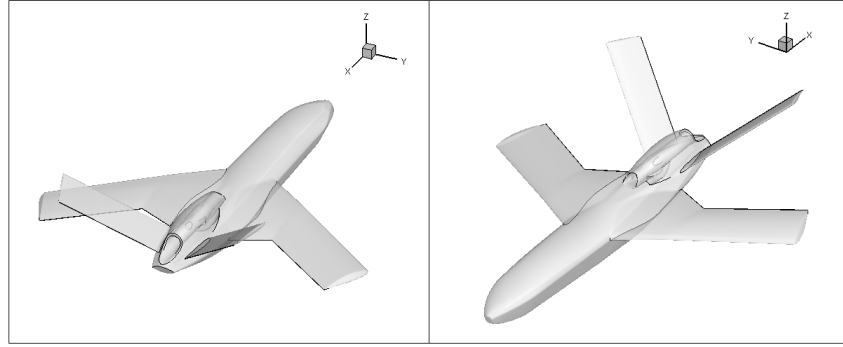


Figure 3-31: Aircraft surrogate model. Visualization of the geometry

3.5.1 Problem Setting

A subsonic flow characterized by a Mach number of 0.25 and a Reynolds number of $4 \cdot 10^6$ is considered. The computational domain is composed by 10^6 points and can be seen in Fig. 3-32. The Alenia UNS3D solver is used for the CFD simulations.

The PODI model is used to build a database for all the aircraft operative configurations therefore the parameter space is large: the angle of attack α varies between 0° and 14° , the sideslip angle β is between 0° and 6° . Taking into account the good performances of the previous subsonic application, a quadtree distribution is exploited to position the snapshots composing the initial high-order simulations. A visualization of the snapshot distribution is presented in Fig. 3-33.

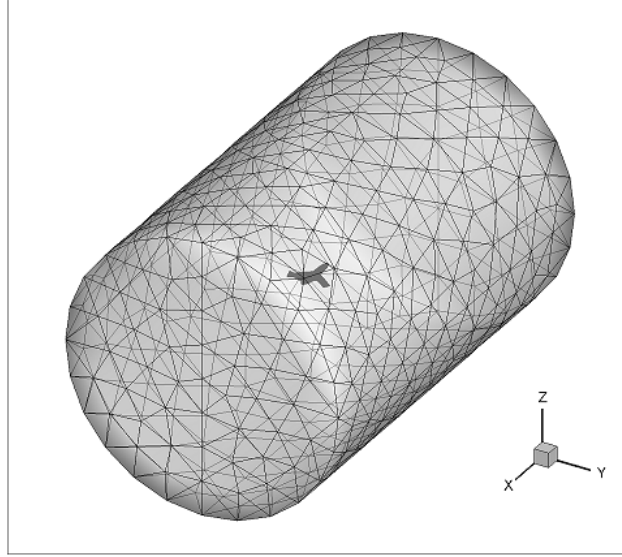


Figure 3-32: Aircraft surrogate model. Visualization of the computational domain

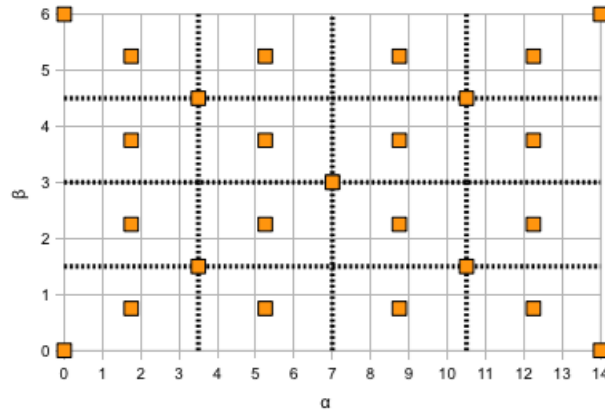


Figure 3-33: Aircraft surrogate model. Snapshot distribution

3.5.2 Results

20 random points are chosen to test the surrogate model. Lift and drag coefficients are computed starting from the fields of the conservation variables generated using

the surrogate model. The results are listed in Table 3.9. The relative error e_{rel} is computed following Eq. (3.6).

Table 3.9: Aircraft surrogate model. Results of the 20 random points

Test Point	α	β	$e_{rel} (c_L)[\%]$	$e_{rel} (c_D)[\%]$
Point 1	4.7	5.8	0.02	11.22
Point 2	3.4	0.3	0.41	11.21
Point 3	7.6	2.2	0.08	1.75
Point 4	10.1	5.2	0.03	1.80
Point 5	12.0	1.4	0.43	0.44
Point 6	3.6	5.3	0.24	9.7
Point 7	8.4	4.4	0.04	0.07
Point 8	10.9	1.4	0.14	0.53
Point 9	12.7	1.4	0.09	1.73
Point 10	1.2	4.9	5.26	7.54
Point 11	4.6	5.7	0.003	8.32
Point 12	4.0	4.2	0.12	1.40
Point 13	12.7	1.7	0.15	1.53
Point 14	8.0	1.0	0.13	0.71
Point 15	6.7	0.7	0.001	2.91
Point 16	8.6	4.9	0.09	0.36
Point 17	1.3	5.1	6.19	4.03
Point 18	12.2	3.8	0.07	0.18
Point 19	3.1	4.0	0.17	0.18
Point 20	7.1	0.5	0.09	4.32

A good agreement with the CFD solution is obtained for the lift coefficient calculation. On the other hand, the drag coefficient comparison is characterized by larger errors. In Fig. 3-34, the points corresponding to higher errors are visualized and it can be remarked that they belong to an 'outer area' of the parameter space.

In the inner area underlined in Fig. 3-35 low errors with respect to the reference solution are obtained considering both lift and drag coefficient computations. For the application of the PODI surrogate model to database generation therefore, this distinction between outer and inner area of the parameter space can be taken into

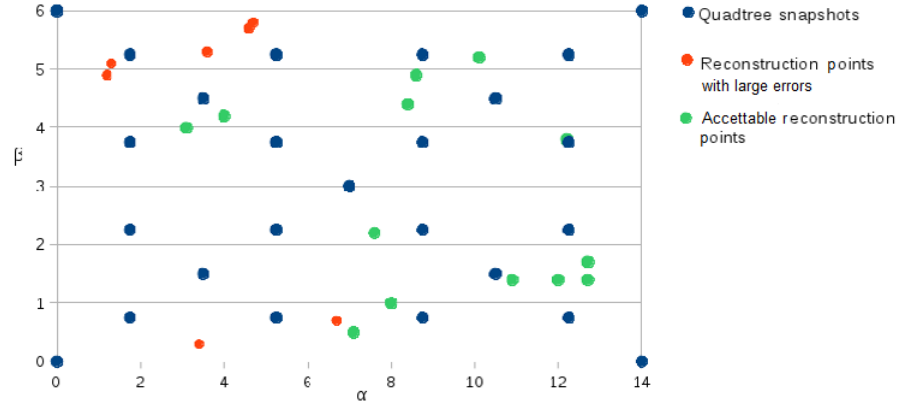


Figure 3-34: Aircraft surrogate model. Large error points distribution

account and the use of the surrogate model inside the internal area is recommended.

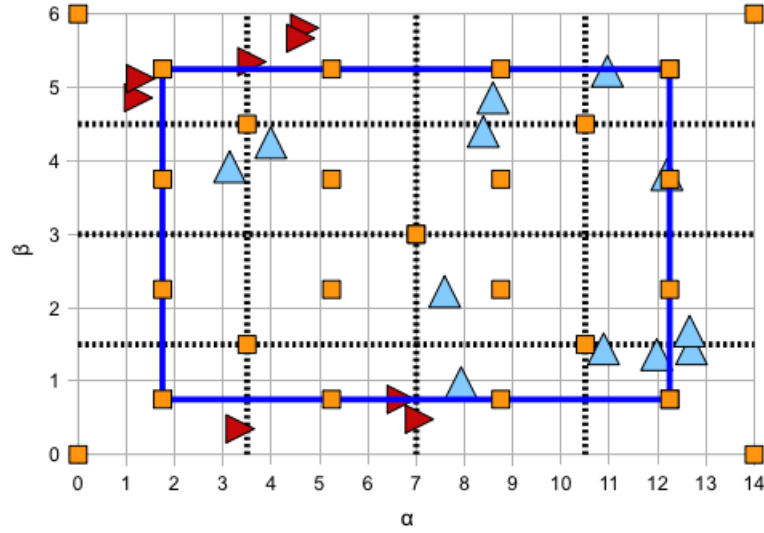


Figure 3-35: Aircraft surrogate model. Inner area of the parameter space

3.6 Optimization of the three-parameter Viermin function

The optimization of the three-parameter Viermin function $f_v(\mathbf{x})$ is performed following the methodology explained in sect. 2.3. The general form of the Viermin function can be written as

$$f_v(\mathbf{x}) = 0.01 \sum_{i=1}^k \left[\left(x_1 + \frac{1}{2} \right)^4 - 30x_i^2 - 20x_i \right]. \quad (3.7)$$

In figure 3-36 the isosurfaces of the three-parameter Viermin function can be visualized.

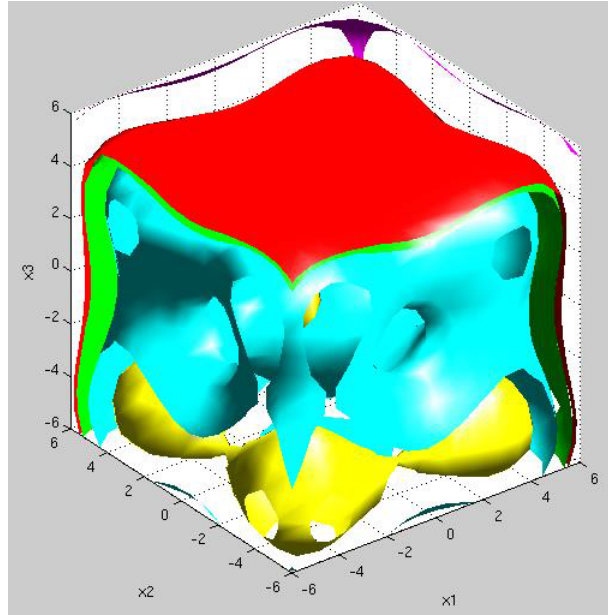


Figure 3-36: Isosurfaces of the three-parameter Viermin function

Matrix V has been constructed using 20000 points and POD was performed. The parameter space is three dimensional and four eigenvectors with four corresponding

eigenvalues were obtained.

The eigenvalues are plotted in fig. 3-37 and they can be seen as a measure of the energy of the system contained in each eigenvector

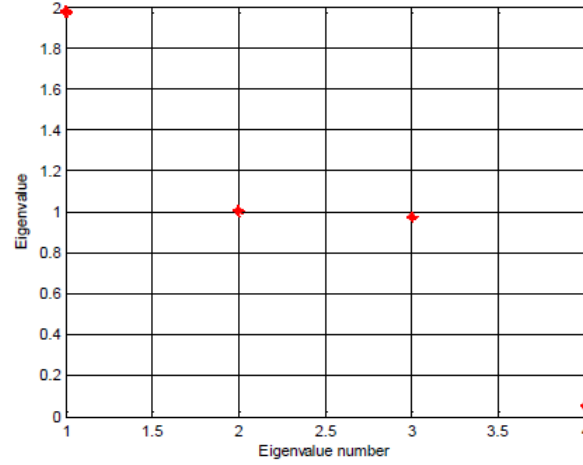


Figure 3-37: The four eigenvalues obtained from POD of matrix V in descending order

In order to perform the minimization of the Viermin function a dimension reduction can be performed and the last eigenvector, corresponding to the less energetic eigenvalue, can be neglected. The projection of matrix V in a reduced-order parameter space can be performed using the $[k+1,k] = [4,3]$ reduction matrix \tilde{A}_r . In this way a description of the problem in a three-dimensional space is obtained. A visualization of the problem in the reduced-order parameter space can be seen in fig. 3-38.

The minimization is performed in the reduced-dimension parameter space using the Matlab function *min* and subsequently the vector containing the optimum point is re-projected in the original space using the pseudoinverse \tilde{A}_r^i of matrix \tilde{A}_r . The results are reported in tab 3.10

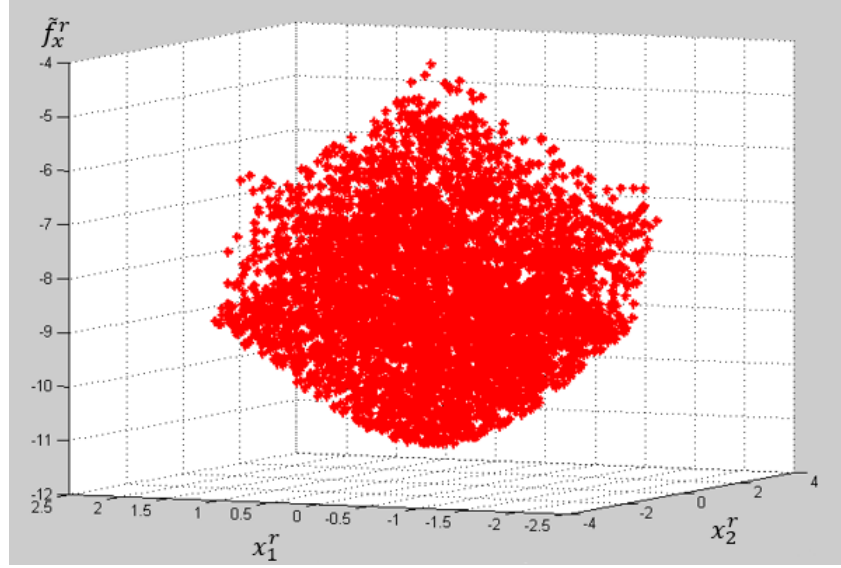


Figure 3-38: Viermin function in a three-dimensional parameter space

Table 3.10: Minimization of the Viermin function. Results

Real minimum			POD minimum			$e_{rel}[\%]$
-7.8491			-7.9254			1
Real optimum point			POD optimum point			e %
-4.4538	-4.4538	-4.4538	-4.6161	-4.6750	-4.5493	5

The errors are computed as

$$e_{rel} = \frac{|p^* - \tilde{p}^*|}{p^*} \cdot 100 \quad (3.8)$$

$$e\% = \max \frac{|x^* - \tilde{x}^*|}{x^*} \cdot 100. \quad (3.9)$$

Chapter 4

Conclusions

In this research work the surrogate model construction process is investigated. The first two chapters introduce the problem framework and a description of some possible distributions of the initial samples needed to build a surrogate model is performed. In chapter two a literature review is given first and subsequently the methods used to construct the surrogate models are described.

Two different approaches are presented and used in the present work: the first one is more classical and a parametric POD is applied exploiting the snapshot method without projection of the model equations in a reduced-order space. The difference with respect to the surrogate models described in literature is that in our case the snapshot set is very small and no truncation of the POD modes is foreseen.

In the second approach, described in sect 2.3, a POD is performed directly in the parameter space in order to reduce the dimensions of the problem. The optimization is performed in this new reduced parameter space and finally the optimum point is re-projected again onto the original space.

Chapter 3 describes numerous applications of the POD surrogate models previously described. In particular approach 1 is used in the backward facing step analysis,

to investigate the two-dimensional flow around a NACA 0012 and a RAE 2822 airfoil, to minimize the drag coefficient of the external shape of a vehicle and to build an aerodynamic database of a three-dimensional aircraft.

In all the applications the surrogate models showed good agreement with the reference CFD solution. Analyses of the influence of the number and distribution of the initial snapshots are presented together with a posteriori error evaluations.

In sec 2.3 the minimization of an analytical function is described using the second approach of a POD reduction of the parameter space. The errors with respect to the true solution are acceptable.

A more detailed error analysis can be performed in further studies, together with some a priori error estimations. An automatic method to establish the optimal number and position of the initial snapshots can be foreseen and the extension of the POD surrogate model to a higher number of parameters would be an interesting research subject.

Bibliography

- [1] A. Sobester A. Forrester and A. Keane. *Engineering Design via Surrogate Modelling - A Practical Guide*. University of Southampton, 2008.
- [2] G. Berkooz, P. Holmes, and J. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Ann. Review Fluid Mech.*, 25:539–575, 1993.
- [3] L. A. Breene. Quadtrees and hypercubes: Grid embedding strategies based on spatial data structure addressing. *The computer journal*, 36(6), 1993.
- [4] D. S. Broomhead and D. Lowe. Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2:321–355, 1988.
- [5] M.D. Buhmann. *Radial Basis Functions: Theory and Implementations*. Cambridge University Press, 2003.
- [6] T. Bui-Thanh, M. Damodaran, and K. Willcox. Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition. *AIAA Journal*, 42(8):1505–1516, 2004.
- [7] B. I. Epureanu. A parametric analysis of reduced order models of viscous flows in turbomachinery. *J. of Fluids and Structures*, 17:971–982, 2003.
- [8] T. Bracconier et al. Toward an adaptive pod/svd surrogate model for aeronautic design. *Computers & Fluids*, 40:195–209, 2011.

- [9] Gerald E. Farin and Dianne Hansford. *The Essentials of CAGD*. A. K. Peters, Ltd., Natick, MA, USA, 1st edition, 2000.
- [10] J. C. Helton and F. J. Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81(1):23–69, July 2003.
- [11] Koziel and Leifsson. *Surrogate-Based Modeling and Optimization Applications in Engineering*. Springer, 2013.
- [12] P. A. Legresley and J. J. Alonso. Airfoil design optimization using reduced order models based on Proper Orthogonal Decomposition. In *AIAA 2000-2545*, Fluids 2000 Conference, Denver, Colorado, June 2000.
- [13] Y. C. Liang, H. P. Lee, S. P. Lim, W. Z. Lin, and K. H. Lee. Proper orthogonal decomposition and its applications - Part I: Theory. *J. Sound and Vibration*, 252(3):527–544, 2002.
- [14] J. L. Lumley. The structure of inhomogeneous turbulent flows. In A. M. Yaglom and V. I. Tatarski, editors, *Atmospheric turbulence and radio propagation*, pages 166–178. Nauka, Moscow, 1967.
- [15] G. Matheron. Le krigeage universel. Technical report, Cahiers du Morphologie Mathématique de Fontainebleau, Ecole des Mines de Paris, 1969.
- [16] M. J. Mifsud, S. T. Shaw, and D. G. MacManus. A high-fidelity low-cost aerodynamic model using proper orthogonal decomposition. *Int. J. Numer. Meth. Fluids*, 63:468–494, 2010.
- [17] Max D. Morris. Factorial sampling planes for preliminary computational experiments. *Technometrics*, 33(2):161–174, May 1991.

- [18] Technische Universität München. Drivaer model. Technical report, <http://www.aer.mw.tum.de/en/research-groups/automotive/drivaer>, 2015.
- [19] R.H. Myers, D. C. Montgomery, and C. M. Anderson-Cook. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. Wiley, 2009.
- [20] S. B. Pope. A more general effective-viscosity hypothesis. *J. Fluid Mech.*, 72:331–340, 1975.
- [21] C. W. Rowley, T. Colonius, and R. M. Murray. Model reduction for compressible flows using POD and Galerkin projection. *Physica D*, 189:115–129, 2004.
- [22] B. Ruck and B. Makiola. Flow separation over the inclined step. *Notes on Numerical Fluid Mechanics*, 40:47–55, 1993.
- [23] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn. Design and analysis of computer experiments. *Statistical Science*, 4(4):409–423, 1989.
- [24] L. Sirovich. Turbulence and the dynamics of coherent structures, part 1: Coherent structures. *Quarterly of Applied Mathematics*, 45(3), 1998.
- [25] B. T. Tan. Proper orthogonal decomposition extensions and their applications in steady aerodynamics. Master thesis, High Performance Computation for Engineered Systems, Singapore-MIT Alliance, 2003.
- [26] L. Tang and W. Shyy. Proper orthogonal decomposition and response surface method for tps/rlv structural design and optimization: X-34 case study. In *AIAA 2005-0839*, 43rd AIAA Aerospace Sciences Meeting, Reno, Nevada, January 2005.